



**Facultad de Biología**

Departamento de Genética

**“Variabilidad genética neutra del  
cromosoma Y de rumiantes domésticos”**

*Lucía Pérez Pardal*

*Santiago de Compostela, 2010*





Los Drs. D. Félix Goyache Goñi y D. Luis José Royo Martín como directores de la Tesis Doctoral titulada “Variabilidad genética neutra del cromosoma Y de rumiantes doméstico” realizada por Dña. Lucía Pérez Pardal en el Departamento de Genética de la Universidad de Santiago de Compostela, informan favorablemente del depósito de la misma, dado que reúne las condiciones necesarias para su defensa.

El Dr. Paulino Martínez Portela como tutor de la Tesis Doctoral titulada “Variabilidad genética neutra del cromosoma Y de rumiantes doméstico” avala la presentación de esta memoria de tesis de Dña. Lucía Pérez Pardal.

Y para que así conste, firman el presente informe en Santiago de Compostela, el 30 de agosto de 2010.

*Fdo.* Félix Goyache Goñi

*Fdo.* Luis José Royo Martín

*Fdo.* Paulino Martínez Portela



*Muchas gracias a mis directores de tesis por todo el trabajo que se han tomado conmigo.*

*Muchas gracias a Paulino Martínez por aceptar ser mi tutor y a quien me lo sugirió.*

*Muchas gracias a todos mis compañeros.*

*Muchas gracias a la gente de Gijón (gracies), a la gente de Minnesota, de Davis (thank you very much) y de Portugal (muita obrigada).*

*Muchas gracias a toda la gente que nos ha enviado muestras.*

*E moitas gracias a mi gente e a miña familia.*





---

<b>a.C.</b>	Antes de Cristo
<b>ADN</b>	Ácido desosiribonucleico
<b>BTAX</b>	Cromosoma X bovino
<b>BTAY</b>	Cromosoma Y bovino
<b>FISH</b>	Hibridación de fluorescencia in situ
<b>HSAX</b>	Cromosoma X humano
<b>HSAY</b>	Cromosoma Y humano
<b>IAM</b>	<i>Infinite Allele Model</i> o modelo de alelos infinitos
<b>IMM</b>	<i>Interspersed multilocus microsatellites</i> o microsatélites <i>multilocus</i>
<b>kb</b>	Kilobase
<b>LINE</b>	<i>Long-interspersed element</i> o elementos nucleares largos
<b>LTR</b>	<i>Long terminal repeats</i> o repetición terminal larga
<b>Mb</b>	Megabase
<b>MSY</b>	Zona específica del macho del cromosoma Y
<b>NRV</b>	Región no recombinante del cromosoma Y
<b>PAB</b>	Limites entre la zona MSY y PAR
<b>PAR</b>	Zona pseudoautosomal del cromosoma Y
<b>pb</b>	Pares de bases
<b>PCR</b>	Reacción en cadena de la polimerasa
<b>QTL</b>	Loci de caracteres cuantitativos
<b>SINE</b>	<i>Short interspersed repetitive elements</i> o elementos nucleares cortos
<b>SMM</b>	<i>Stepwise Mutation Model</i> o modelo mutacional paso a paso
<b>SNP</b>	Polimorfismo de nucleótido simple
<b>sp.</b>	Especie
<b>STR</b>	<i>Short Tandem Repeat</i> o repeticiones en tandém o microsatélites
<b>Yp</b>	Brazo corto del cromosoma Y







*Índice*



<b>1. Justificación del trabajo .....</b>	<b>15</b>
1.1 Presentación y planteamiento .....	17
1.2 Objetivos .....	21
1.3 Financiación .....	22
1.4 Organización de la presente Memoria .....	23
<b>2.Revisión bibliográfica.....</b>	<b>25</b>
2.1 La domesticación.....	27
2.1.1 Características de la domesticación.....	27
2.1.2 Fases de la domesticación .....	28
2.1.3 Origen de la domesticación .....	29
2.1.4 Implicaciones genéticas de la domesticación .....	33
2.2 El ganado bovino: <i>Bos taurus</i> y <i>Bos indicu</i> .....	34
2.2.1 Sistemática .....	35
2.2.2 Características generales .....	36
2.2.2.1 El antecesor común: El Uro .....	37
2.2.3 Caracterización del ganado bovino actual con estudios genéticos por vía materna y autosomal.....	41
2.3 El cromosoma Y .....	46
2.3.1 La naturaleza del cromosoma Y .....	46
2.3.2 Evolución y origen de los cromosomas sexuales .....	47
2.3.2.1 Mecanismos de adquisición de genes del cromosoma Y de mamíferos .....	49
2.3.2.2 Factores de degradación o especialización de genes del cromosoma Y .....	51
2.3.3 Estructura del cromosoma Y en mamíferos .....	55
2.3.3.1 El cromosoma Y humano.....	55
2.3.3.1.1 Clasificación de las secuencias eucromáticas según su origen.....	58
2.3.3.1.2 Clasificación de las secuencias eucromáticas según sus características.....	62
2.3.3.2 El cromosoma Y bovino. ....	62

2.3.3.2.1 Mapa del cromosoma Y bovino.....	65
2.3.3.2.2 Microsatélites descritos en el cromosoma Y bovino .....	67
2.3.3.2.3 Análisis de la variabilidad genética del cromosoma Y bovino.....	69
<b>3. Artículos .....</b>	<b>73</b>
3.1 Artículo 1: Female segregation patterns of the putative Y- chromosome-specific microsatellite markers INRA124 and INRA126 do not support their use for cattle population studies. ....	75
3.2 Artículo 2: Y-specific microsatellites reveal an African subfamily in taurine ( <i>Bos taurus</i> ) cattle.....	81
3.3 Artículo 3: Multiple paternal origins of domestic cattle revealed by Y-specific interspersed multilocus microsatellites .....	91
3.4 Artículo 4: Genetic structure of the bovine Y- specific microsatellite UMN0103 reflects the genetic history of the species. ....	101
<b>4.Discusión.....</b>	<b>109</b>
4.1 Contextualización.....	111
4.2 Estructura de la información .....	115
4.3 Criterios de selección de marcadores específicos del cromosoma Y en ganado bovino.....	116
4.3.1 Criterios generales .....	117
4.3.1.1 Especificidad de macho .....	117
4.3.1.2 Polimorfismo .....	118
4.3.1.3 Segregación .....	119
4.3.2 Criterios adicionales .....	120
4.3.2.1 Genotipado automatizable .....	120
4.3.2.2 Repetibilidad.....	121
4.4 Aplicación de los criterios de selección de marcadores .....	122
4.4.1 Primera selección de marcadores.....	122
4.4.1.1 Rendimiento filogenético de los microsatelites seleccionados .....	124
4.4.2 Segunda selección de marcadores. Desarrollo de nuevas herramientas.....	127
4.4.2.1 Rendimiento filogenético de los IMM.....	130
4.5 Caracterización del microsatélite UMN0103 .....	135
4.6 Repercusión, implicaciones y perspectivas futuras de este trabajo .....	138

---

<b>5. Conclusiones.....</b>	<b>145</b>
<b>6. Resúmenes .....</b>	<b>149</b>
6.1 Sumary .....	151
6.2 Resumen.....	154
6.3 Resumen .....	157
<b>7. Bibliografía.....</b>	<b>159</b>
<b>8. Anexos: .....</b>	<b>177</b>
8.1 Material suplementario del artículo 1.....	179
8.2 Material suplementario del artículo 2.....	183
8.3 Material suplementario del artículo 3.....	188
8.4 Material suplementario del artículo 4.....	193





*Justificación del Trabajo*





## 1.1 Presentación y planteamiento

En la presente Memoria se exponen un conjunto de trabajos cuya temática se resume en el título de la misma: “Variabilidad genética neutra del cromosoma Y de rumiantes domésticos”. En concreto estos trabajos se centran en a) la identificación de marcadores específicos del cromosoma Y bovino y b) la demostración de su utilidad para el estudio de la historia genética de la especie bovina por la línea paterna. Esta memoria incluye aspectos metodológicos, de interés en genética animal, y aspectos de interés multidisciplinar que resultan de interés potencial para un variado espectro de científicos, biotecnólogos, conservacionistas y responsables del manejo y conservación de poblaciones animales domésticas. El estudio de los patrones de dispersión y los procesos de introgresión por la vía masculina en el ganado bovino doméstico (*Bos taurus* y *B. indicus*) es, indudablemente, un marcador potencial de la historia de las poblaciones humanas. Los resultados del trabajo deberían contribuir a mejorar el conocimiento sobre la estructura y diversidad del cromosoma Y bovino, pero también deberían aportar evidencias sobre los procesos genéticos y antropológicos asociados a la domesticación y estructura genética de la especie bovina. Si las sociedades humanas primitivas trataron de manera diferente a machos y hembras, esta diferente presión de selección debe verse reflejada en la diversidad de los marcadores del cromosoma Y respecto de la de marcadores de origen materno y autosómico.

En los últimos 20 años, el estudio genético de los procesos de domesticación, así como de las rutas de expansión y procesos de introgresión entre diferentes poblaciones ha experimentado un gran desarrollo. El análisis de secuencias de ADN mitocondrial (Loftus *et al.*, 1994a; Mannen *et al.*, 2004; Beja-Pereira *et al.*, 2006; Edwards *et al.*, 2007a, b; Achilli *et al.*, 2008; Chen *et al.*, 2010) junto con el hallazgo de ciertas evidencias arqueológicas y, en menor medida, los análisis de marcadores autosómicos de tipo microsatélite (MacHugh *et al.*, 1997) dibujan con

gran precisión una panorámica de lo que sucedió hace aproximadamente 10000 años cuando fue domesticado el ancestro del ganado bovino actual, el uro (*B. primigenius*). Sin embargo, desde el punto de vista genético existen lagunas que conviene aclarar. La más importante consiste en conocer qué sucedió por la vía masculina ¿Se produjo un solo proceso de domesticación de sementales o este proceso se produjo en varios momentos a lo largo de la historia?, ¿Cuál es la variabilidad del cromosoma Y del ganado bovino actual?, ¿Cómo influyeron posibles diferencias de manejo entre las líneas paternas y maternas en el acervo genético de la especie y la filogeografía?, ¿Cuál es la fiabilidad de los trabajos desarrollados hasta el momento en este campo?.

La historia de las líneas paternas de la especie bovina permanece todavía en la incertidumbre. Para paliar esta carencia de conocimiento son de especial interés los marcadores localizados en la región específica masculina del cromosoma Y (MSY) ya que se heredan como un clon de padre a hijo sin que se produzca recombinación entre ellos. Esta característica los diferencia de los marcadores localizados en la región PAR del cromosoma Y, donde sí se produce recombinación con el cromosoma X. En el caso de la especie humana el análisis de la diversidad del cromosoma Y ha demostrado su utilidad para el establecimiento de relaciones filogenéticas y evolutivas (Kayser *et al.*, 2000; Rosser *et al.*, 2000; Hurles y Jobling, 2001), así como su importancia en genética forense y estudios de enfermedades (Odelberg *et al.*, 1988; Evett *et al.*, 1991; Wiegand *et al.*, 1992; Kayser *et al.*, 1997; Jobling y Tyler-Smith, 2000; Ali y Hasnain, 2002). Sin embargo, pese a todo el conocimiento acumulado en la especie humana, en lo relativo a poblaciones animales, y especialmente en rumiantes, apenas existen estudios filogenéticos basados en la diversidad genética del cromosoma Y (Van Hooft *et al.*, 2002; Tosi *et al.*, 2003; Götherström *et al.*, 2005; Pidancier *et al.*, 2006). La escasez de estudios realizados en este cromosoma se explica parcialmente porque se trata de un cromosoma que solo recombina con su

homólogo, el X, en una pequeña porción (la zona pseudoautosomal o PAR) y además está plagado de secuencias repetitivas, cuestiones que dificultan su análisis. La complejidad de su estudio se refleja en que cada vez que se inicia un proyecto de secuenciación del genoma de una especie, el individuo elegido es siempre una hembra. El genoma bovino publicado por The Bovine Hapmap Consortium en abril de 2009 no resulta una excepción a esta regla. Después de seis años de trabajo, un grupo de 300 investigadores de 25 países logró descifrar completamente la secuencia del genoma bovino con una excepción: el cromosoma Y. Este hecho hace más atractivo aún trabajar en este campo.

El objetivo de este trabajo es ayudar a mitigar este déficit de información existente sobre la diversidad genética del cromosoma Y del ganado bovino. Para ello, se propuso el estudio del cromosoma Y bovino en una línea de trabajo diferente de la marcada por trabajos pioneros en este campo, en el que se centran en la búsqueda y genotipado de SNP en secuencias de genes localizados en la zona MSY bovina (Götherström *et al.*, 2005). Desde este trabajo se propone hacer filogenia con microsatélites específicos de la zona MSY del cromosoma Y. Los microsatélites presentan una serie de características generales que les confieren un grado de preferencia frente a otros marcadores:

- a) Se trata de secuencias que se encuentran en un número alto a lo largo del genoma y repartidas de forma más o menos homogénea, una cada 10 ó 20 kb (Edwards *et al.*, 1991, Edwards *et al.*, 1992; Bowcock *et al.*, 1994; Forbes *et al.*, 1995).
- b) A priori tienen un efecto neutro frente a la selección.
- c) Son fácilmente amplificables a través de la PCR. La metodología necesaria para su estudio es sencilla, rápida y de bajo coste económico.

d) El resultado de la amplificación es visible mediante geles de poliacrilamida o mediante secuenciadores automáticos, y la caracterización en pares de bases de los alelos es sencilla, ayudada de programas informáticos, de modo que los resultados entre laboratorios son fácilmente intercambiables.

e) Poseen un elevado polimorfismo como consecuencia de una mayor tasa de mutación ( $10^{-3}$  a  $10^{-5}$  mutaciones por gameto) (Edwards *et al.*, 1992; Bowcock *et al.*, 1994; Forbes *et al.*, 1995).

f) Se encuentran bastante conservados desde el punto de vista evolutivo de modo que unos microsatélites descubiertos en una especie pueden ser amplificados con éxito en otra especie del mismo género, aunque el nivel de polimorfismo encontrado no sea el mismo.

g) Son marcadores codominantes.

h) Alelos poco frecuentes pueden estar asociados con caracteres de interés.

i) Se ha estimado que son necesarios aproximadamente 5 SNP para equiparar la información que proporciona un microsatélite (Wilson y Sorant, 2000).

Debe tenerse en cuenta que las características arriba mencionadas hacen referencia a microsatélites autosómicos y que los microsatélites situados en la zona macho específica (o MSY) del cromosoma Y no cumplen todas estas características como por ejemplo que son marcadores dominantes. Las características generales que presentan los microsatélites, junto con la batería de

marcadores microsatélite descrita para el cromosoma Y bovino (Liu *et al.*, 2003) han hecho que estos marcadores fuesen los idóneos para realizar este estudio.

El primer hito propuesto en esta memoria consistió en realizar una labor de criba y selección de los microsatélites descritos hasta el momento como específicos de macho en el cromosoma Y de bovinos. Para ello, hemos utilizado como información previa los trabajos de Matthews *et al.*, (1991), Bishop *et al.*, (1994), Vaiman *et al.*, (1994), Mezzelani *et al.*, (1995), Kappes *et al.*, (1997), Liu *et al.*, (2003), todos ellos desarrollados en ganado bovino, y el reciente trabajo de Meadows *et al.*, (2006) desarrollado en ovino. Los marcadores seleccionados se testaron en una amplia variedad de poblaciones bovinas pertenecientes a las especies *B. taurus* y *B. indicus* para comprobar su utilidad en estudios filogenéticos. El muestreo realizado, que comprende la mayor parte de razas bovinas europeas, así como poblaciones de cebú africano y asiático, ha constituido una parte muy importante del esfuerzo realizado.

## **1.2 Objetivos**

Una vez establecido el interés del trabajo a desarrollar, nos planteamos una serie de objetivos, que se pueden clasificar como de carácter metodológico y de carácter multidisciplinar:

Como objetivos de carácter metodológico se plantearon los siguientes:

- a) Identificar marcadores microsatélite polimórficos específicos de la región MSY del cromosoma Y bovino.
- b) Establecer los protocolos que permitan testar la utilidad de los marcadores identificados para la evaluación de la variabilidad genética de las líneas del ganado bovino.
- c) Desarrollar protocolos de diagnóstico útiles para el conocimiento de la variabilidad genética existente en el cromosoma Y de rumiantes domésticos.

Como objetivos de carácter multidisciplinar se propusieron los siguientes:

- d) Establecer el origen (monofilético o polifilético) de las líneas paternas de ganado bovino doméstico, *B. taurus* y *B. indicus*.
- e) Determinar, por la vía macho, la existencia de una estructura genética en el ganado bovino doméstico.
- f) Establecer las posibles rutas de expansión y la existencia de procesos de introgresión de las diferentes familias genéticas de ganado bovino en áreas geográficas diferentes.
- g) Estimar tiempos de divergencia entre las familias haplotípicas identificadas en el cromosoma Y bovino.

Además como objetivo no previsto en el planteamiento inicial del trabajo:

- h) se ha determinado la utilidad y demostrado la naturaleza repetitiva de las nuevas herramientas desarrolladas en esta Memoria.

### 1.3 Financiación

Parece claro que los objetivos generales planteados presentan interés general. En ese sentido, cabe destacar que el trabajo que se resume en esta Memoria ha recibido el apoyo del Ministerio de Ciencia e Innovación mediante la concesión, en concurrencia competitiva, de los siguientes proyectos de investigación: CGL2005-03761/BOS (Identificación de polimorfismos en el cromosoma y de pequeños rumiantes para su utilización en filogenia) y CGL2008-03949/BOS (Uso combinado de marcadores polimórficos de evolución rápida y lenta en la filogenia del cromosoma Y de pequeños rumiantes). Por parte del Gobierno del Principado de Asturias (PCTI2009) se concedió el proyecto IB-09-144 (Polimorfismos de los microsatélites *multilocus* en la filogenia del cromosoma Y de pequeños rumiantes).

Asimismo, la beca FPI (BES-2006-13545) asociada al proyecto CGL2005-03761/BOS ha permitido la elaboración de esta Memoria. Parte del trabajo realizado se ha desarrollado en laboratorios distintos al del Área de Genética y Reproducción Animal del SERIDA. Estancias de investigación financiadas el Ministerio de Ciencia e Innovación permitieron trabajar con los grupos del Dr.

Ponce de León en la Universidad de Minnesota, de la Dra. Cecilia Penedo en la Universidad de Davis, California, y del grupo del Dr. Albano Beja-Pereira del CIBIO de la Universidad de Oporto. En esas estancias se han podido aprender y contrastar técnicas de análisis que han permitido llevar a cabo el trabajo propuesto. Durante la estancia realizada en la Universidad de Oporto se desarrolló el trabajo que permitió la redacción del artículo científico titulado “Y-specific microsatellites reveal an African subfamily in taurine (*B. taurus*) cattle”. Esta estancia de 3 meses permite presentar esta Memoria para optar a la mención especial de doctorado europeo.

#### **1.4 Organización de la presente Memoria**

Esta Memoria se presenta en la modalidad de compendio de publicaciones. El esquema utilizado consta de **justificación del trabajo**, parte en la que nos encontramos, donde se presentan el problema a resolver y los objetivos planteados. Una **revisión bibliográfica** sirve para contextualizar el problema planteado en este documento. La siguiente sección se corresponde con las **publicaciones científicas** a que ha dado lugar el trabajo realizado y se continúa con una **discusión** general en la que se valoran conjuntamente las publicaciones presentadas. En el siguiente capítulo se presentan las **conclusiones** obtenidas en este trabajo. A continuación los **resúmenes** del trabajo que son presentados en inglés, castellano y gallego. Finalizando se encuentra el capítulo dedicado a la **bibliografía** utilizada en la presente memoria y un capítulo dedicado a los **anexos** que incluyen el material suplementario a los artículos. Una copia de esta memoria se adjunta en formato electrónico (disco compacto).







## *Revisión bibliográfica*



Esta sección se ha desglosado en 3 bloques, La domesticación, El ganado bovino y El cromosoma Y. La finalidad de esta división es contextualizar los tres campos que incumben a esta memoria.

## **2.1 La domesticación**

La domesticación puede definirse como un proceso mediante el cual una población animal se adapta al hombre y a una situación de cautividad a través de una serie de modificaciones genéticas que suceden en el curso de generaciones (Price, 1984).

### **2.1.1 Características de la domesticación**

La domesticación es un proceso de adaptación a un ambiente nuevo en el que se establece una necesidad de una adecuación gradual al ser humano y a condiciones ambientales nuevas derivadas del confinamiento y el control reproductivo de los animales domesticados (Price, 1984). Este proceso precisa de largos períodos de tiempo y el paso de numerosas generaciones, para que los cambios producidos (de comportamiento, morfológicos, fisiológicos u otros) se fijen genéticamente. Las poblaciones domésticas se han seleccionado por una serie de características como docilidad, producción de leche, carne, trabajo o locomoción. Todo esto ha llevado a un proceso de reducción de tamaño de los animales, un acortamiento del rostro y reducción de los dientes, una variación de las características morfológicas según el clima y cambios en el pelaje (Zeder *et al.*, 2006).

Todos los cambios producidos en los animales hacen que se considere como animal doméstico el que cumple la siguiente serie de normas (Zeder *et al.*, 2006; Price, 1984):

- a) Integración a la cultura humana.

- b) Mantenimiento bajo el control del hombre para un propósito definido.
- c) Dependencia de éste para su supervivencia.
- e) Reproducción bajo condiciones artificiales controladas por el ser humano.
- f) Modificación en razas especializadas por la selección artificial.

En cuanto a la elección de animales a domesticar, hay una serie de características que han influido para que sean unos animales y no otros los elegidos. Estas características son: la dieta, la tasa de crecimiento, la reproducción en cautividad, baja agresividad, no tendencia al pánico y la estructura social. Debido a estas cuestiones sólo un pequeño número de especies han terminado siendo objeto de domesticación (Zeder *et al.*, 2006).

### **2.1.2 Fases de la domesticación**

La domesticación es un proceso continuo en el que se reconocen una serie de fases. Según Zeuner (1963), existen cinco etapas fundamentales dentro del proceso de domesticación:

- \* Primera etapa: se caracteriza por que el control que el hombre ejerce sobre el animal es muy reducido. Son frecuentes los cruces de las formas mantenidas en cautividad con las formas salvajes originarias.
- \* Segunda etapa: el hombre comienza a controlar la reproducción de los animales y a seleccionarlos para reducir sus dimensiones y aumentar las características de docilidad. La finalidad de este proceso es un mejor manejo de los animales. Es muy importante evitar el cruce con las formas salvajes para mantener y conseguir fijar las características deseadas.

- \* Tercera etapa: el hombre comienza a demostrar un interés creciente hacia la producción de carne, y se da cuenta de la utilidad que supone el aumento de las dimensiones de los animales de cría. Por ello se vuelven a cruzar las formas domésticas, más pequeñas, con las formas salvajes, más grandes, con el objetivo de un aumento en el tamaño pero manteniendo las características de docilidad previamente seleccionadas.
- \* Cuarta etapa: el interés se centra en los productos de origen animal y en la creciente capacidad del hombre para controlar a los animales de producción. Mediante un largo trabajo de selección, esta etapa concluye con la creación de razas especializadas con diferentes aptitudes productivas.
- \* Quinta etapa: nuevamente, el hombre evita los cruces de la forma salvaje con las razas domésticas especializadas. Se realiza una actividad de control numérico de la población salvaje, que en muchos casos acaba con el exterminio de las formas salvajes o su asimilación dentro de las formas domésticas.

Hart (1985) considera que actualmente nos encontraríamos en una Sexta etapa del proceso de domesticación. La etología y genética de los animales de producción se habrían visto modificadas hasta tal punto que los animales habrían perdido la capacidad de sobrevivir y de reproducirse sin la intervención del hombre. Sin embargo, muchas de las características perdidas pueden ser readquiridas en procesos de readaptación a la vida salvaje.

### **2.1.3. Origen de la domesticación**

La domesticación de animales y plantas es uno de los sucesos más importantes de la historia de la humanidad. Se trata de un proceso tan importante que marcó el final de una época, el paleolítico, y el inicio de otra, el neolítico. Se produjo un

cambio en la sociedad pasando de una sociedad nómada cazadora-recolectora a una sociedad agroganadera, lo que permitió asegurar una fuente de alimentación continua, reduciendo la incertidumbre de tener que depender de la caza o de la recolección silvestre, y a su vez, posibilitó la formación de asentamientos fijos.

La mayor parte de las especies agroganaderas que hoy conocemos se domesticaron hace unos 7.000-10.000 a.C. (Bruford *et al.*, 2003). El descubrimiento de la agricultura comenzó con el cultivo de los cereales silvestres en el arco de montañas y tierras altas de Oriente Próximo, sin perjuicio de la domesticación de diversas especies vegetales en diferentes lugares de América, Europa y Asia. Desde Oriente Próximo, esta tecnología se extendería a todo el mundo (Salamini *et al.*, 2002). La domesticación animal es un proceso complejo que se llevó a cabo en varias partes del mundo y en distintas ocasiones a lo largo del tiempo. En la actualidad se reconocen tres áreas principales de domesticación animal en el mundo: Oriente Próximo, este de Asia y la cordillera de los Andes en Sudamérica (Brudford *et al.*, 2003; Zeder *et al.*, 2006; Kadwell *et al.*, 2001). En el Oriente Próximo (el Creciente Fértil y su margen oriental hacia la region del valle del Indo) se domesticó el cerdo (*Sus scrofa*), la oveja (*Ovis aries*), la cabra (*Capra hircus*) y la vaca (*B. taurus*) (Larsson, 2005, 2007; Luikart *et al.*, 2001; Loftus *et al.*, 1994 a; Meadows *et al.*, 2007). Se conocen procesos locales de domesticación animal en el este de Asia en el cerdo (*S. scrofa*), el búfalo (*Bubalus bubalis*), la oveja (*O. aries*), la cabra (*C. hircus*), el cebú (*B. indicus*), el yak (*B. grunniens*) y el gaur (*B. gaurus*) (Zeder *et al.*, 2006). En la cordillera de los Andes entre 6.000-7.000 a.C se domesticó la llama (*Lama glama*) y la alpaca (*Vicugna pacos*) a partir del guanaco (*L. guanicoe*) y la vicuña (*V. vicugna*) (Kadwell *et al.*, 2001; Bruford *et al.*, 2003), Tabla 1.

**Tabla 1.** Cronología de la domesticación

Nombre común	Especie	Fecha	Lugar	Referencia
Cabra	<i>Capra sp.</i>	10.000 a.C	Oriente Próximo	Legge (1996)
Oveja	<i>Ovis sp.</i>	9.000 a.C	Oriente Próximo	Legge (1996)
Cerdo	<i>Sus sp.</i>	9.000 a.C	Oriente Próximo, China	Clutton-Brock (1979)
Ganado	<i>Bos sp.</i>	8.000 a.C	Oriente Próximo, India	Clutton-Brock (1989)
Caballo	<i>Equus caballus</i>	6.000 a.C	Estepas Eurasiáticas	Clutton-Brock (1992)
Búfalo acuático	<i>Babalus babalis</i>	6.000 a.C	China	Olsen (1993)
Camélidos	<i>Llama sp.</i>	6.000 a.C	Cordillera de los Andes	Novoa y Wheeler (1984)
Burro	<i>Equus asinus</i>	5.000 a.C	NE África	Clutton-Brock (1992)
Camellos	<i>Camelus sp.</i>	5.000 a.C	Arabia (Dromedary) Asia (Bactrian)	Köhler-Rollefson (1996)
Yak	<i>Bos grunniens</i>	>2.000 a.C	Tibet	Olsen (1990)
Reno	<i>Rangifer tarandus</i>	<2.000 a.C	Norte de Eurasia	Aikio (1989)

Con el descubrimiento de la agricultura se propició el cambio de vida del hombre de recolector nómada a agricultor sedentario, dándose las condiciones ideales para que el hombre comenzase la domesticación de los pequeños rumiantes. Hace aproximadamente 10.000 a.C. se domesticó la cabra en la cadena montañosa que cubre desde el este de Anatolia hasta el norte y el centro de Zagros en Irán (Zeder *et al.*, 2000; Smith, 1995; Luikart *et al.*, 2001; Naderi *et al.*, 2007, 2008). Por la misma época y en la misma zona geográfica se domesticó la oveja (Meadow, 1993; Chessa *et al.*, 2009; Zeder *et al.*, 2006). Ambos animales pertenecen al orden Artiodactyla, familia Bovidae y tribu Caprine. Son animales con gran capacidad para adaptarse a distintas condiciones ambientales. La cabra (*C. hircus*) se domesticó a partir del bezoar (*C. aegagrus*) en un proceso complejo que pudo incluir dos procesos de domesticación separados en el tiempo que afectaron a la misma metapoblación de cabra salvaje (Naderi *et al.*, 2007, 2008). El posible origen polifilético de la cabra doméstica, que incluiría la domesticación del markhor (*C. falconeri*) parece descartable ya que evidencias arqueológicas y morfológicas no apuntan en ese sentido (Manceau *et al.*, 1999, Mannen *et al.*,

2001, Takada *et al.*, 1997). Este hecho se ve reforzado además por análisis filogenéticos usando el Citocromo b y el Cromosoma Y, en los que *C. aegagrus* y *C. hircus* clusterizan juntos (Zeder *et al.*, 2006). De forma parecida, la oveja (*O. aries*) se pudo domesticar en dos fases separadas en el tiempo. La segunda de ellas podría ser resultado de la búsqueda de producciones especializadas como la lana o la leche (Chessa *et al.*, 2009). Parece claro que el origen de la oveja sería el muflón asiático (*O. orientalis*) (Zeder *et al.*, 2006; Meadows *et al.*, 2007). Existe consenso científico en que el muflón europeo (*O. musimon*), que fue considerado durante algún tiempo el ancestro más probable de la oveja doméstica, es una reliquia asilvestrada de las primeras ovejas domésticas introducidas en Europa hace 7.000 a.C. (Clutton-Brock, 1999; Meadows *et al.*, 2007; Chessa *et al.*, 2009). La carencia de registro fósil de muflones en el neolítico europeo apoya esta idea que, además, está sustentada por estudios filogenéticos en los que muestras de muflón clusterizan entre muestras de ovejas domésticas (Zeder *et al.*, 2006).

La domesticación de pequeños rumiantes como los ovinos y los caprinos favoreció a su vez la domesticación de animales de mayor envergadura, como los bovinos (*B. taurus* y *B. indicus*), ya que con la leche de los primeros se pueden alimentar las crías de los segundos, haciendo más manejable al animal ya desde joven (Zeder *et al.*, 2006). La domesticación de los bovinos además de proporcionar abonos, pieles, carne y utilizarse como animal de tiro y transporte también influyó en el proceso de desmonte de los bosques y generación de pasto. A todas estas ventajas hay que añadir que de ellos se podía obtener una fuente de alimentación como es la leche y todos sus derivados (MacHugh *et al.*, 1997). La domesticación de un animal de grandes dimensiones también se vio ayudado de los primeros animales domesticados, los perros, que ayudarían a cuidar de los rebaños y manadas (Vilà *et al.*, 1997).



No todos los procesos de domesticación animal resultan igual de nítidos. Un ejemplo de este hecho es la domesticación del caballo (*Equus caballus*). El inicio de la domesticación de esta especie pudo tener lugar hacia el 6.000 a.C. en las estepas eurasiáticas. Este proceso ha sido complejo y comprende varias áreas del mundo (Ucrania, las estepas rusas y Kazajstán) en distintos periodos de tiempo que podrían ir del 6000 a. C al 4000 a. C. (Vilà *et al.*, 2001; Jansen *et al.*, 2002; Zeder *et al.*, 2006). El inmediato progenitor del caballo no ha sido claramente definido, aunque se apunta a que podrían ser caballos salvajes de Eurasia. Tampoco está suficientemente esclarecido el papel de África en la domesticación animal, sólo existe consenso científico en el reconocimiento del burro (*E. asinus*) como especie procedente de ese continente (Beja-Pereira *et al.*, 2004) aunque muchos autores, como se comentará mas adelante, sugieren una posible domesticación africana de *B. taurus* (por ejemplo: Bradley y Magee, 2006). Esto hace pensar en un posible centro de domesticación animal probablemente en la zona del cuerno de África que engloba a Somalia, Djbuti, Eritrea y Etiopía (Beja-Pereira, comunicación personal).

#### **2.1.4 Implicaciones genéticas de la domesticación**

El origen de las sociedades agrícolas y ganaderas se ha asociado a dos hechos: el uso de una parte muy reducida de la biodiversidad existente en cada una de las áreas en que estas sociedades fraguaron, y la adaptación de las especies elegidas a nuevas condiciones favorables al uso humano (domesticación). Pero en general la adaptación a las condiciones de domesticación se opone a los mecanismos naturales de adaptación y está asociada a cambios morfológicos y fisiológicos. En muchos casos estos cambios hacen que sea difícil reconocer la especie salvaje de la que surgió la domesticada correspondiente. La domesticación por lo tanto es un proceso que reduce la variabilidad genética en la población domesticada respecto del ancestro salvaje. Este proceso se llevó a cabo a partir de un pequeño número de individuos capturados de la población salvaje. Al aislar a la población domesticada de la población salvaje se provoca un efecto fundador que conlleva una importante

pérdida de la diversidad genética. Este hecho, unido a la selección de características deseables, ha uniformizado genéticamente las poblaciones domésticas. Con la domesticación se están perdiendo alelos y características que poseían los animales salvajes. La conservación del ancestro salvaje resulta fundamental como reserva genética para casos de emergencia y para futuros procesos de selección, ya que los ancestros salvajes portan variedades genéticas asociados a resistencia a plagas y enfermedades, genes que pueden mejorar la productividad, dar mayor valor nutritivo, buen sabor y muchas otras propiedades, además de que la variabilidad genética asegura una mejor respuesta de una población o especie a condiciones ambientales variables y cambiantes. Una población genéticamente variable presenta en general una respuesta mucho más plástica frente a diferentes condiciones ambientales, mientras que en las poblaciones uniformes todos los individuos responden de forma similar o en el mismo grado. El peligro reside por tanto en que ante determinadas condiciones (climáticas, presencia de patógenos, etc.) todos los animales domesticados son igualmente susceptibles pudiendo, en la situación más extrema, llegar a extinguirse. En el caso del ganado bovino no se conserva el ancestro, hecho que hace, sin duda, más importante el estudio de la diversidad para evaluar y conservar correctamente la variabilidad existente, sobre todo teniendo en cuenta el proceso de erosión genética sufrido en el último siglo: la uniformidad ha sustituido a la diversidad a nivel mundial. En los últimos 100 años se han perdido, aproximadamente, unas 1000 razas y en la actualidad unas 2000 están en peligro (FAO, 2000).

## **2.2 El ganado bovino: *Bos taurus* y *Bos indicus***

El ganado bovino actual incluye dos especies, *B. taurus* y *B. indicus*, así como los cruces generados entre estas dos especies.

### 2.2.1 Sistemática

Estas dos especies se clasifican de la siguiente manera:

Especie: *B. taurus*; *B. indicus*

Género: *Bos*

Familia: *Bovidae*

Orden: *Artiodactyla*

Clase: *Mammalia*

Filo: *Chordata*

Reino: *Animalia*

Los bóvidos pertenecen al orden zoológico de los artiodáctilos e incluyen animales diversos como la gacela (*Gazella* spp.), el antílope (*Tetracerus quadricornis*), el búfalo africano (*Syncerus caffer*) y la cabra montés (*C. pyrenaica*), y especies domésticas como la oveja (*O. aries*), la cabra (*C. hircus*) y la vaca (*B. taurus*). Esta familia, la Bovidae, abarca 137 especies comprendidas en 45 géneros. Sus miembros cubren un extenso rango de climas y hábitats diferentes, extendiéndose a través de África, gran parte de Europa, de Asia y América. En la actualidad, África es el continente con mayor diversidad de especies de bóvidos.

Dentro de la familia Bovidae se encuentra la subfamilia Bovinae que está dividida en tres tribus: Tragelaphini o Strepscerotini, Boselaphini y Bovini. Estos últimos divididos en búfalos (*Bubalus* sp. y *Synarus* sp.) y no búfalos (*Bos* sp. y *Bison* sp.). Los géneros *Bos* y *Bison* tienen un cariotipo de  $2n = 60$  con autosomas acrocéntricos o telocéntricos (Gallagher y Womack, 1992), a diferencia del búfalo africano (*Synarus* sp.) con  $2n = 52$  y el búfalo acuático (*Bubalus* sp.) con  $2n = 50$  ó  $48$ , ya sea de río o pantano, respectivamente. Estas diferencias en el número de cromosomas, probablemente explicaría el porqué de la ausencia de cruces entre estas especies. La tribu Bovini (subfamilia Bovinae, familia Bovidae) contiene el mayor número de especies que han sido domesticadas.

### 2.2.2 Características generales

En la actualidad, se distinguen dos grandes grupos de bovinos domésticos: el ganado vacuno (*B. taurus*) y el cebú (*B. indicus*) que a su vez se clasifican según su localización geográfica, historia y morfología (Epstein y Mason, 1984; Felius, 1995). En Europa los individuos de *B. taurus* se organizan en razas mientras que en Asia y África se habla de poblaciones debido a la carencia de registros genealógicos y programas de selección. La principal distinción entre estos animales se encuentra en la posesión o no de giba (*B. indicus* y *B. taurus*, respectivamente, Figura 1). *B. taurus* predomina en los climas templados mientras que *B. indicus* habita terrenos áridos o semiáridos de África y Asia. Además de por la presencia de giba, el cebú difiere del ganado vacuno en presentar una cabeza más alargada y una mayor resistencia a garrapatas y parásitos gastrointestinales. El ganado vacuno presenta una mayor resistencia a la picadura de la mosca tse-tse en regiones de África donde este invertebrado está presente (Epstein y Mason, 1984; Marshal, 1989). A pesar de sus diferencias, el cebú y el ganado vacuno, presentan capacidad de adaptación en ambientes diferentes. Tanto cebú como toros presentan el mismo número de cromosomas ( $2n = 60$ ) lo que hace que se encuentren algunas razas intermedias, genéticamente de origen mixto, que en África se denomina *sanga* (Felius, 1995; Frisch *et al.*, 1997; MacHugh *et al.*, 1997).

**Figura 1.** Fotografías de un *B. taurus* y un *B. indicus* en los que se puede ver la marcada diferencia entre estas dos especies.



#### 2.2.2.1 El antecesor común. El uro

El antecesor común de los bóvidos actuales es el uro (*B. primigenius*). Es frecuente ver a este animal esquematizado en numerosas pinturas rupestres junto a otras especies, demostrando que en el pasado fueron muy abundantes (Figura 2). El Uro fue un bovino de cabeza grande, cuernos muy largos que sobrepasaban los 1,20 metros de largo, presentaba cuello corto, tronco amplio, región dorso-lumbar recta y llena, tórax profundo y vientre voluminoso. La domesticación de un animal de estas dimensiones (alzada a la cruz de 2 metros y una longitud escápulo-isquial de 2,50 m), fue un paso formidable para el hombre siendo uno de los puntos definitivos en el desarrollo humano. Además de obtener sus productos (leche, carne, pieles y estiércol), el hombre podría utilizar como animal de trabajo, sirviendo de instrumento para el desarrollo de la agricultura y la ganadería (Payne, 1991). En Europa se ha constatado la existencia de uros hasta épocas recientes. Debido a la pérdida de su hábitat y a la presión de caza, la población salvaje de uros no ha sobrevivido hasta la actualidad. En el siglo XIII era un animal restringido al este de Europa, y en el siglo XVII existía un rebaño en el bosque de Jaktorow (Polonia), a cuyo último ejemplar, hembra, se dio caza en 1627 (Mertens, 1906).

**Figura 2.** Detalle de un uro encontrado en las pinturas rupestres de la Cueva del Pudial (Ladruñan, Sistema Ibérico Aragonés).



El uro habría evolucionado a partir del *B. acutifrons* asiático hace unos 5-1.8 millones de años durante el Plioceno superior en la India (Zeuner, 1963; Van Vuure, 2005). Los uros migraron al Oriente Medio a través de Asia y habrían llegado a Europa hace unos 250.000 años (Van Vuure, 2005). Su distribución abarcaba el continente asiático, norte de África y el continente europeo, incluyendo las islas británicas y la península escandinava. Basándose en evidencias arqueológicas, algunos arqueozoólogos proponen separar el *B. primigenius* en 3 subespecies. Se encontraron fósiles que presentan diferencias morfológicas en los cuernos y diferencias de tamaño por lo que proponen una subespecie euroasiática (*B. primigenius primigenius*), otra presente en el sureste asiático (*B. primigenius namadicus*) y una tercera subespecie ligada al norte de África (*B. primigenius opisthonomus*) (Clutton-brock, 1989). El cebú de la India está relacionado con uros que divergieron de los uros del Oriente Próximo hace al menos 200.000 años. El ganado africano sería descendiente de uros más relacionados con los uros del Oriente Próximo. Los uros del Oriente Próximo y los de África habrían divergido hace 25.000 años, unos 15.000 años antes de la domesticación. Los uros presentes en Mongolia, norte de China, Corea y Japón podrían haber divergido de los uros del Oriente Próximo hace unos 35.000 años (Mannen *et al.*, 2004). La expansión

del bovino desde sus centros de domesticación en Asia se habría producido entre 8.000 y 10.000 años a.C. (Payne, 1991; Loftus *et al.*, 1994 a y b).

Parece claro que el antecesor de *B. taurus* y *B. indicus* fue el *B. primigenius*, lo que no es tan diáfano fue como se produjo su domesticación. Algunos autores (Epstein y Mason, 1984; Payne, 1991), han sugerido que el ganado vacuno y el cebú provienen de una única domesticación en el Oriente Próximo. Sin embargo, análisis de ADN mitocondrial han probado que el ganado bovino moderno proviene, al menos, de dos procesos diferentes de domesticación (Bradley *et al.*, 1996; Loftus *et al.*, 1994 a, b; Troy *et al.*, 2001; Helmer *et al.*, 2005). Este hecho confirma informaciones arqueológicas previas (Zeuner, 1963; Grigson, 1980; Meadow, 1993) que sugerían que la domesticación ocurrió en dos o más puntos geográficos y temporales diferentes, manifestando que la morfología del cebú se derivaría de una población asiática de *B. primigenius* diferente de la que habría dado lugar al ganado vacuno. Posteriormente, estudios con microsatélites autosómicos (Kim *et al.*, 2002) sobre distintos animales en el Noroeste de Asia y un vasto estudio del ADN mitocondrial del ganado bovino de Asia, muestran patrones de diversidad genética en *B. indicus* que apoyan que la domesticación del cebú ocurrió en las cuencas de los ríos Indo y Ganges (Chen *et al.*, 2010). Domesticaciones independientes podrían haber sucedido además en África (Bradley *et al.*, 1998, Hanotte *et al.*, 2002; Zeder *et al.*, 2006) y en el sudeste asiático (Mannen *et al.*, 2004). Análisis de ADN mitocondrial de uros los sitúan en el macro-haplogrupo P, que divergió antes de la domesticación del macro-haplogrupo T pero después de la separación de la rama del macro-grupo I (Troy *et al.*, 2001; Bailey *et al.*, 1996; Edwards *et al.*, 2007; Stock *et al.*, 2009), Figura 3.

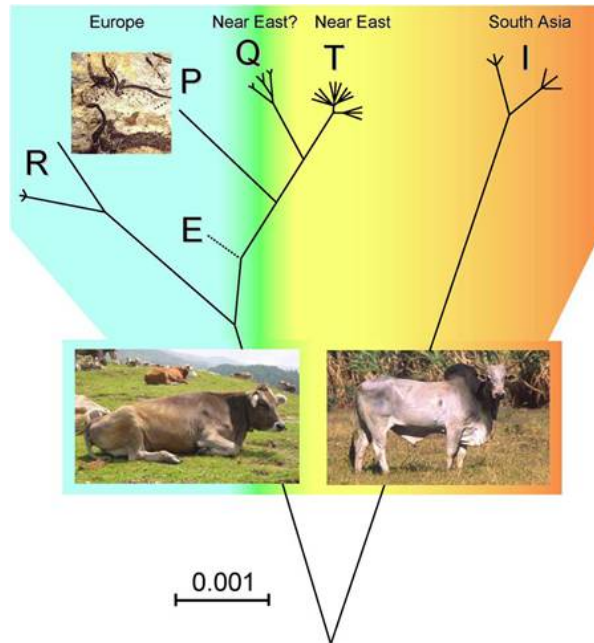
Existe un gran debate científico sobre la posibilidad de que poblaciones locales de *B. primigenius* se intrograsaran en las poblaciones de bovino doméstico,

influyendo en mayor o menor medida en el acervo genético del ganado bovino actual.

El hecho de que el ganado bovino europeo pertenezca casi exclusivamente al macro-haplogrupo T, junto con la ausencia del haplogrupo P, sustentaría la idea de que hembras de uro no se incorporaron al pool genético domesticado (Troy *et al.*, 2001; Edwards *et al.*, 2007; Magee *et al.*, 2002). En cuanto a la vía paterna los trabajos de Bollonguino *et al.*, (2008) y Svenson y Göthrström (2009) apoyarían esta hipótesis. Los trabajos de Achilli *et al.*, (2008, 2009) y Stock *et al.*, (2009), con ADN mitocondrial, aportan una visión diferente de estos hechos ya que encontraron en animales modernos italianos la presencia de individuos pertenecientes al haplogrupo P y otros animales que presentaban los haplogrupos Q y R, que son haplogrupos que se cree que son originarios de uros. Además se ha detectado el haplogrupo T en uros europeos (Beja-Pereira *et al.*, 2006). Estos trabajos sugieren la introgresión de uros en las líneas maternas. Estudios complementarios en el cromosoma Y también sugieren introgresión por vía paterna (Göthrström *et al.*, 2005).



**Figura 3.** Árbol filogenético más parsimónico construido a partir de ADN mitocondrial de Bovinos (Achilli *et al.*, 2009). Este árbol sin raíz incluye todas las muestras disponibles de ADN mitocondrial completo de los meta-haplogrupos P, Q y R descritos en *B. primigenius*. La rama se basa en el análisis parcial de la región control y la región codificante y también está relacionado este meta-haplogrupo con *B. primigenius*. I y T se han descrito en *B. indicus* y *B. taurus*, respectivamente. La pintura paleolítica del uro pertenece a la cueva de Lascaux, en el sur de Francia.



### 2.2.3 Caracterización del ganado actual con estudios genéticos por vía materna y autosomal

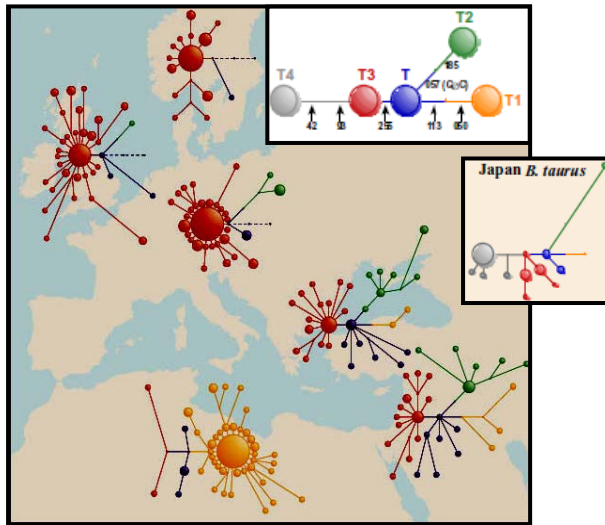
En los últimos años el número de trabajos de caracterización genética del ganado bovino se ha incrementado notablemente. Los análisis de secuencias de ADN mitocondrial (Bradley *et al.*, 1996), análisis de microsatélites nucleares y estudios de polimorfismos bioquímicos (Manwell y Baker, 1980 a, b; MacHugh *et al.*, 1997) permiten deducir las posibles migraciones de *B. taurus* y *B. indicus* desde su centro de domesticación, así como conocer la diversidad genética existente en la actualidad dentro de especie y entre especies, además de ver como se distribuye esa variabilidad a nivel geográfico. La variabilidad genética en

poblaciones de *B. taurus* se estimó entre el 4 y el 8% con respecto al total de los mamíferos utilizando polimorfismos proteicos (Baccus *et al.*, 1983). Se puede considerar que la pérdida de variabilidad genética del ganado vacuno se debe al efecto fundador al que se encontró sometido el ganado actual. Por otro lado, los procesos de incorporación de nuevos uros dentro del *pool* doméstico pueden haber sido una repetición de las mismas poblaciones, provocando que la variación se mantuviera muy constante, con la problemática añadida de la pérdida de la diversidad genética inducida por la domesticación (Fries y Ruvinsky, 1999). La variabilidad genética también se estudió utilizando microsatélites, que constataron la existencia de una sustancial diversidad dentro y entre poblaciones bovinas (MacHugh *et al.*, 1997). Sin embargo, estos datos son difíciles de comparar con otras especies, debido a que parten de la predisposición de una alta variabilidad en la especie origen.

En cuanto al estudio de las líneas maternas, diversos estudios se han realizado permitiendo describir los haplogrupos maternos. Estudios complementarios de la región control del ADN mitocondrial, D-loop, en 34 razas actuales de Europa, Oriente Próximo y África (Troy *et al.*, 2001), permitieron aislar 152 haplotipos distintos que difieren en 77 sitios polimórficos. Se definieron cuatro haplogrupos de los cuales el más predominante fue el haplogrupo conocido como T3, seguido por el haplogrupo T1 (transiciones en las posiciones 16.050, 16.113 y 16.255, número posicional que ocupa el nucleótido en la secuencia del ADN mitocondrial), el T2 (transiciones 16.185 y 16.255; transversión 16.057), y, por último, el haplogrupo T (transición en la posición 16.255). El haplogrupo T3 corresponde en su mayor parte a los individuos pertenecientes a las zonas europeas (Franja noroeste, Gran Bretaña y Europa Continental), el T1 presenta la mayoría de sus haplotipos representados en África. El T2 está repartido por la zona de Europa Continental, Anatolia y Oriente Medio mientras que el haplogrupo T se

correspondería a una secuencia central, que se encuentra en Anatolia y Oriente Medio, que puede ser la más parecida a la secuencia origen (Figura 4).

**Figura 4.** Distribución geográfica de los haplogrupos T, T1, T2, T3 y T4 de *B. taurus* (en formato Network), figura adaptada de Troy et al., 2001. En esta figura se puede apreciar que los haplogrupos T1, T2 y T3 presentan divergencia en forma de estrella.



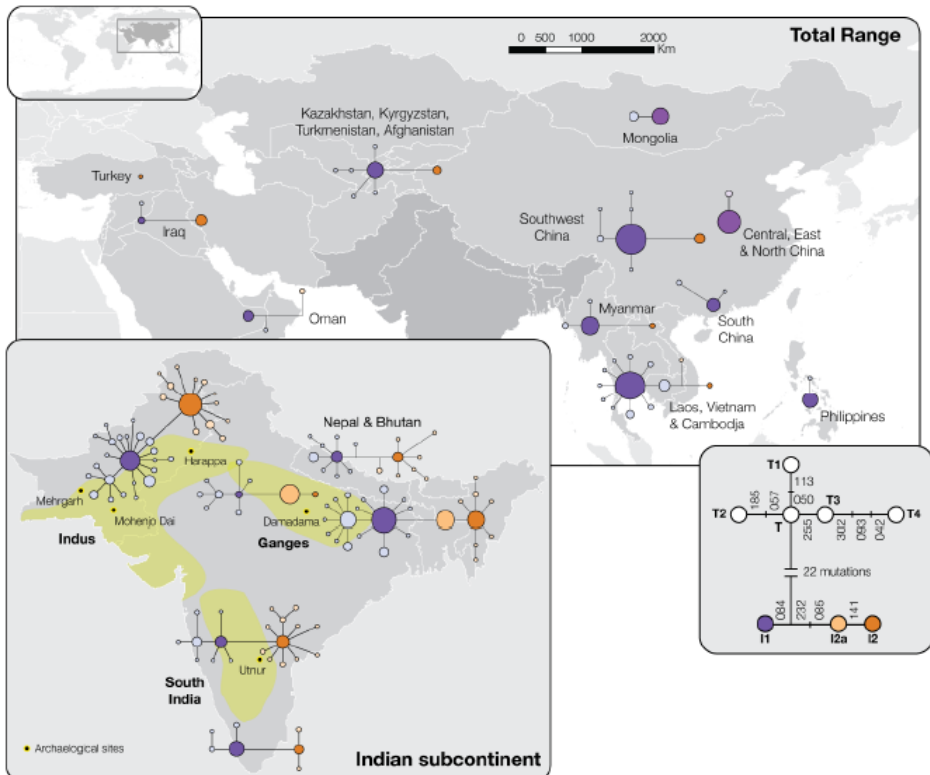
Posteriormente, Mannen *et al.*, (2004) describieron otro nuevo haplogrupo T4, con tres nuevas transiciones (16.042, 16.093 y 16.302) derivadas del haplogrupo T3, cuyos representantes fueron descritos por primera vez en el ganado de Japón, sugiriendo un origen mitocondrial independiente en el noreste asiático. Achilli *et al.*, (2008) identifican un nuevo haplogrupo, el T5, además de encontrar en animales actuales el haplogrupo Q, relacionado este último con los uros, lo que indicaría introgresiones recientes de estos animales. Los haplogrupos T1, T2 y T3 presentan una divergencia con forma de estrella, lo que indica una rápida expansión. La divergencia dentro del grupo T, es mucho menor de la que presentan el ADN mitocondrial de humanos, lo que sugiere un cuello de botella en la reciente historia de *B. taurus* (Achilli *et al.*, 2008). Respecto al mismo análisis realizado con poblaciones de *B. indicus* se hallaron dos haplogrupos: I1 e I2 (Lai *et al.*, 2006;

Baig *et al.*, 2005; Achilli *et al.*, 2008). Esos dos haplogrupos están bien representados en India (Baig *et al.*, 2005; Chen *et al.*, 2010), existiendo dos teorías que expliquen este hecho: una que sugiere que se domesticaron dos poblaciones distintas de hembras proto-cebuinas (Baig *et al.*, 2005) o, una segunda teoría más probable, que hace referencia a la existencia de un único proceso de domesticación de ganado I1 en el Valle del Indo (aproximadamente 4.000 años a.C.) con una introgresión secundaria de hembras I2 (aproximadamente 1.500 años a.C.) en los rebaños de cebúes previamente domesticados (Chen *et al.*, 2010). La divergencia entre los grupos T se estima en unos 14.100–44.500 (Lai *et al.*, 2006) mientras que la divergencia entre *B. taurus* y *B. indicus* respecto al ancestro común sucedió hace cientos de años, indicando domesticaciones independientes (Lai *et al.*, 2006; Achilli *et al.*, 2008). Los patrones de diversidad y la distribución geográfica del cebú I1 y I2 no han sido bien estudiadas debido a la falta de muestras de la India y mayor parte de China (Lai *et al.*, 2006, Baig *et al.*, 2006). El haplogrupo I1 se encontró en la India y China mientras que el I2 parece estar restringido a animales de la India (Chen *et al.*, 2010) (Figura 5).

Otra aproximación que permite aportar nuevas evidencias respecto a la domesticación de *B. taurus* y la variabilidad pérdida debido a este proceso es el análisis de ADN antiguo. Secuencias obtenidas de ADN procedente de huesos de *B. primigenius* británicos se compararon con secuencias de procedencia moderna y se comprobó que estas secuencias eran genéticamente distintas, con lo que se concluyó que no había existido domesticación en Europa (Bailey *et al.*, 1996; Troy *et al.*, 2001). De este hecho se dedujo que las líneas maternas de los individuos europeos se originaron en el oriente próximo, por lo que toda la diversidad del ADN mitocondrial presente en el ganado actual se ha acumulado desde los tiempos neolíticos (Bollongino *et al.*, 2006). La domesticación del ganado es un proceso que ocurrió hace 10.000 años, donde sólo un pequeño número de uros, de los que descienden los 1.300 millones de cabezas de ganado que hay en la actualidad,

fueron domesticados. La domesticación es una fuerza evolutiva que ha afectado a la especie de distintos modos, produciendo cuellos de botella, deriva y flujos génicos de unas poblaciones a otras. Por el momento hay información de como ocurrió el proceso de la domesticación por la vía materna, aunque en ciertos aspectos aún se plantean discrepancias, por ello es importante hacer un nuevo aporte de información obtenida por otras vías. El estudio del cromosoma Y aportaría luz al proceso acontecido por la vía paterna. De todos modos se debe tener en cuenta que además de estar influido por la domesticación, el cromosoma Y posee unas características propias que afectan a su contenido informativo.

**Figura 5.** Distribución geográfica de los haplogrupos 11 e 12 de *B. indicus* (en formato Network) en el continente asiático (Chen et al., 2010). Se hace énfasis en los posibles centros de domesticación dentro de la India que se representan en verde (Valle del Indo, del Ganges y sur de la India).



## 2.3 El cromosoma Y

El cromosoma Y es uno de los cromosomas sexuales, sólo presente en los individuos macho. Está compuesto por dos zonas bien diferenciadas, la MSY o específica de macho y la PAR o pseudoautosomal. Respecto de su secuencia, el cromosoma Y es único entre los cromosomas ya que tiene largas porciones especulares, o palindrómicas, no encontrándose palíndromos de esta escala y grado de precisión en ningún otro lugar del genoma.

### 2.3.1 La naturaleza del cromosoma Y

Durante mucho tiempo el cromosoma Y se consideró como un cromosoma que no contenía información genética a excepción del gen SRY o de determinación del sexo. La mayor parte de su secuencia parecía no tener sentido y estaba plagado del denominado ADN basura o *Junk DNA* por lo que se creía que este cromosoma era un vertedero de ADN. Sin embargo trabajos recientes en humanos (Skaletsky *et al.*, 2003; Jobling *et al.*, 2003) han demostrado que en la región específica masculina del cromosoma Y (MSY) están presentes más de 20 genes o familias de genes.

Una de las características principales del cromosoma Y es que no tiene un cromosoma homólogo con el que recombinar durante la división celular como mecanismo de reparación de los genes que hayan podido sufrir mutaciones deletéreas. Asumiendo que el mismo proceso de mutación actúa en todos los cromosomas, la diversidad esperable en la secuencia del cromosoma Y es más baja que la del resto del genoma nuclear (Hammer *et al.*, 1995; Thomson *et al.*, 2000). Hay que considerar que por cada cuatro copias de cada autosoma hay tres cromosomas X pero una sola copia del cromosoma Y. Por consiguiente, el tamaño efectivo de población del cromosoma Y es un cuarto de la de cualquier autosoma, un tercio de la del cromosoma X y similar a la del ADN mitocondrial haploide. Este cromosoma es mucho más susceptible a la deriva genética que otros cromosomas, lo que puede provocar cambios aleatorios en las frecuencias de

haplotipos de una generación a la siguiente. La deriva acelera la diferenciación entre grupos de cromosomas Y en diferentes poblaciones, una propiedad bastante útil para investigar la historia de las poblaciones. Sin embargo, debido a la deriva, las frecuencias de haplotipos pueden cambiar rápidamente a lo largo del tiempo. Así, las huellas genéticas que procesos de mestizaje entre poblaciones o cambios demográficos pasados han podido dejar en el cromosoma Y de las poblaciones estudiadas podrían ser inestables, dando información que podría inducir a errores.

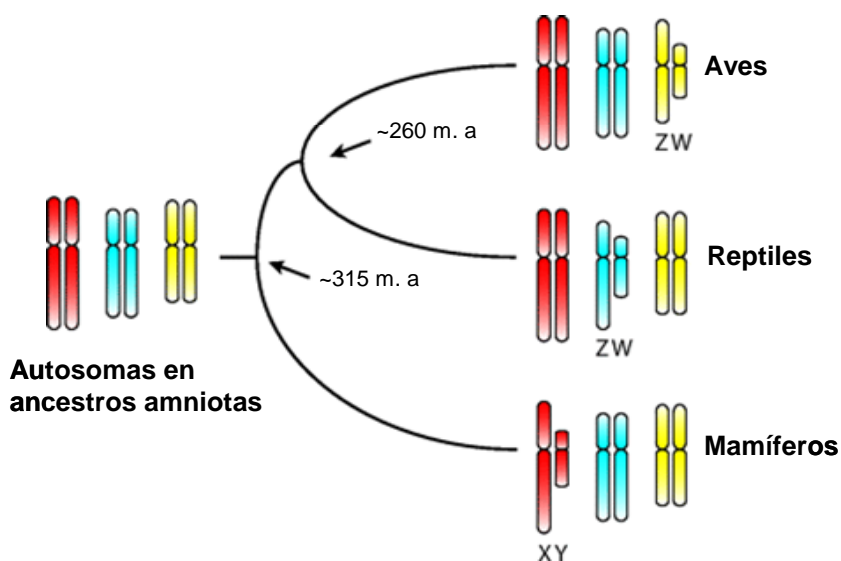
### **2.3.2 Evolución y origen de los cromosomas sexuales**

Los cromosomas sexuales surgieron hace millones de años a partir de cromosomas autosómicos (Figura 6). Hay varios tipos de sistemas cromosómicos sexuales y distintos mecanismos para determinar el sexo en los animales. Antes de la aparición de los cromosomas sexuales el sexo se determinaba por factores ambientales o por otra serie de mecanismos genéticos. En mamíferos, salvo contadas excepciones, el sexo está determinado por la presencia de un cromosoma Y y un cromosoma X en machos y la presencia de dos cromosomas X en hembras. En otros amniotas, como es el caso de aves y reptiles, el sexo está determinado por una variedad de mecanismos genéticos o ambientales, encontrándose especies que combinan ambos mecanismos (Quinn *et al.*, 2007). Los cromosomas sexuales se denominan X e Y cuando el sexo heterogamético es el masculino (hembras XX y machos XY) y se utilizan las denominaciones ZW cuando el sexo heterogamético es el femenino (hembras ZW y machos ZZ), como ocurre en reptiles o aves. El sistema XY y los cromosomas sexuales WZ son aparentemente similares. Los cromosomas X y Z en el sexo homogamético están por lo general bien conservados, son de un tamaño considerable y ricos en genes, mientras que la mayoría de los cromosomas sexo específicos (Y y W) son normalmente pequeños, heterocromáticos y poseen una pequeña cantidad de genes. Se conocen excepciones a esta regla general como son los cromosomas sexuales gigantes de

*Microtus agrestis*, cuyos cromosomas sexuales presentan un mayor tamaño que los autosomas (Ohno, 1967).

Mapectos comparativos muestran que los cromosomas sexuales de mamíferos (XY), aves (ZW) y reptiles (ZW) son enteramente no homólogos, lo que implica que derivan de parejas de autosomas diferentes (Fridolfsson *et al.*, 1998; Nanda *et al.*, 2000; Graves y Shetty 2001; Kohn *et al.*, 2004; Matsubara *et al.*, 2006). El gen que determina el sexo, encargado de iniciar el desarrollo de los testículos es también diferente. En muchos vertebrados el gen encargado de determinar el sexo es el SRY. En aves es un gen situado en el cromosoma Z, posiblemente el gen *DMRT1* (Raymond *et al.*, 1999). En los animales invertebrados no existe equivalente de este gen.

**Figura 6.** Modelo de evolución independiente de los cromosomas sexuales de aves, reptiles y mamíferos a partir de cromosomas autosomales (Vallender *et al.*, 2006). Amniotas ancestrales no poseen cromosomas sexuales y el sexo se determina, probablemente, por temperatura. Los cromosomas sexuales evolucionarían a partir de autosomas en tres ocasiones diferentes, siendo el origen de los cromosomas sexuales de aves, reptiles y mamíferos tres parejas de autosomas distintos. El sistema ZZ:ZW surgiría dos veces, (una en aves y otra en reptiles) mientras que el sistema XX:XY surgiría una sola vez en mamíferos.



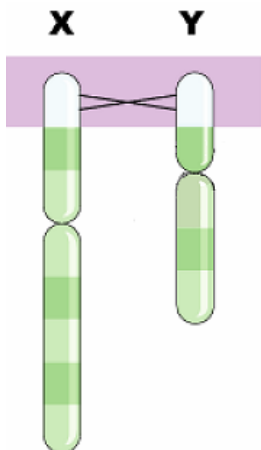


Un caso intermedio entre los sistemas XY y ZW lo presenta el ornitorrinco (*Ornithorhynchus anatinus*), cuyo sistema es XY, poseyendo las hembras 5 parejas XX y los machos cinco parejas XY. Lo más sorprendente es que pese a presentar sistema XY el contenido génico muestra homologías con el sistema ZW de aves presentando el gen *DMRT1* en uno de los cromosoma X (Veyrunes *et al.*, 2008). Este complejo sistema podría ser el nexo de unión entre los sistemas XY y ZW que se creía que habían evolucionado de forma independiente (Grützner *et al.*, 2004; Ezaz *et al.*, 2006) y las semejanzas mostradas por los distintos sistemas de cromosomas sexuales serían el resultado de la convergencia de las fuerzas evolutivas en diferentes linajes de vertebrados. Los cromosomas sexuales de mamíferos derivan de una pareja de cromosomas autosómicos que dejaron de recombinar entre ellos hace 166 millones de años después de la separación de los monotremas. Los cromosomas X e Y de los terios (marsupiales y placentarios) son por lo tanto 145 millones de años más jóvenes de lo que se había estimado (Veyrunes *et al.*, 2008). La adquisición de un gen de determinación del sexo por uno de los miembros de la pareja de autosomas y la acumulación de alelos sexo específicos en esa zona, seguida de la supresión de la recombinación entre los nuevos proto-cromosomas sexuales, tiene como resultado la aparición de una zona vulnerable a la degradación por mutación o delección en los cromosomas sexuales heterogaméticos (Y o W) (Charlesworth y Charlesworth, 2000; Charlesworth y Charlesworth 2005; Fraser y Fraser, 2005; Steinemann y Steinemann, 2005; Graves, 2006).

#### **2.3.2.1 Mecanismos de adquisición de genes del cromosoma Y de mamíferos**

El cromosoma Y de mamíferos presenta por norma general dos zonas bien diferenciadas, la zona pseudoautosomal y la zona específica de macho, también denominada MSY (Figura 7).

**Figura 7.** Esquema de los cromosomas sexuales X e Y. En el recuadro rosa se muestra la zona PAR (zona en la que se pueden producir eventos de recombinación), el resto se corresponde con la zona MSY.



El cromosoma Y de mamíferos solo sobrecruza con el cromosoma X por la zona PAR, siendo esta la única zona donde es posible la recombinación con el cromosoma X. Análisis comparativos de los cromosomas sexuales de mamíferos sugieren que el cromosoma Y adquiere genes a partir de la región pseudoautosomal. Estos genes serán luego incorporados a la región MSY mediante dos procesos: la inversión y la translocación intracromosomal (Ellison *et al.*, 1996; Lahn y Page, 1999).

La inversión es un tipo de aberración cromosómica en la que un segmento de un cromosoma gira 180° sobre sí mismo. La inversión puede comprender unas pocas bases dentro de un gen o largas regiones de ADN que contienen varios genes. El segmento invertido puede ser corto o muy largo y puede incluir o no al centrómero. Una inversión no implica pérdida de información genética sino simplemente una reordenación lineal de la secuencia de genes. Este proceso requiere dos roturas y la posterior reinserción del segmento invertido. Se forma un lazo cromosómico antes de la rotura, los nuevos extremos creados se aproximan y se reúnen. Cuando se produce la inversión en grandes regiones del cromosoma Y,

este proceso hace que se cambie el orden de las secuencias respecto del cromosoma X con la consecuencia de que se produce el cese del solapamiento entre la mayor porción de estos dos cromosomas.

La translocación intracromosomal es el desplazamiento de un segmento de un cromosoma a un nuevo lugar en el mismo cromosoma (en este caso el cromosoma Y). Los genes localizados en ese segmento cambian su posición dentro de ese cromosoma. El resultado es la incorporación de esos genes desde la zona PAR a la zona MSY.

Independientemente de cual sea el mecanismo por el que se incorpora un gen a la región MSY, se produce el cese del entrecruzamiento y la recombinación entre el cromosoma X e Y. Los genes de la zona MSY se especializarán en funciones masculinas o sufrirán una degradación gradual debido a distintos procesos (Graves, 1995) que se describen en el siguiente punto.

#### **2.3.2.2 Factores de degradación o especialización de genes del cromosoma Y**

Hay una serie de factores responsables de la degradación o especialización de las funciones de los genes presentes en el cromosoma Y. A continuación se enumeran y explican las que presentan una mayor repercusión.

- **Deleción:** este proceso produce la pérdida de material genético de un cromosoma que puede ir desde la pérdida de un solo nucleótido (deleción puntual) hasta la pérdida de grandes regiones evidenciables mediante técnicas citogenéticas.
- **Muller *rachet* o trinquete de Muller:** es un proceso descrito en poblaciones asexuales por el que se acumulan en el genoma mutaciones perjudiciales de forma irreversible. En ausencia de recombinación, las mutaciones deletéreas se acumulan rápidamente.

Este proceso es típico del cromosoma Y a excepción de la zona PAR (Charlesworth y Charlesworth, 2000; Figura 8).

- Efecto *Hitchhiking* (o *Selective sweep* o barrido selectivo): este mecanismo causa la selección positiva a favor de las mutaciones que aporten alguna ventaja evolutiva. Las mutaciones neutras ligadas a estas mutaciones ventajosas acaban fijándose en la población con la consecuencia de pérdida de diversidad (Rice, 1987). El impacto de este proceso depende de la tasa de recombinación y del coeficiente de selección. A menor tasa de recombinación y mayor coeficiente de selección, mayor será la región afectada por el efecto *Hitchhiking*. En el cromosoma Y el 95% de los *loci* presentes están ligados por lo que un gen ventajoso (favorable) que regule alguna función relacionada con la condición masculina, como la espermatogénesis, podría fijarse en la población y eliminar toda la variación de la zona MSY (Roldan y Gomendio, 1999; Wyckoff *et al.*, 2000). Este proceso podría fijar un cromosoma Y dentro de una especie y promovería diferencias entre especies (Figura 8).
- Selección *Background*: es un efecto de la selección negativa donde mutaciones deletéreas van a ser eliminadas de la población y con ellas serán eliminadas las mutaciones neutras ligadas a estas. Se produce, por lo tanto, una disminución de la variabilidad alrededor del sitio seleccionado. Como ocurre con el efecto *Hitchhiking* el impacto de esta selección negativa depende de la tasa de recombinación y del coeficiente de selección. La Selección *Background* no cambia las frecuencias alélicas con el mismo alcance que el efecto *Hitchhiking*, lo que hace que estos dos procesos se puedan distinguir uno del otro (Charlesworth y Charlesworth, 1995). En una región no recombinante,

alelos ligeramente deletéreos y ligeramente ventajosos pueden sobrevivir ligados. En ausencia de una fuerte mutación beneficiosa, una mutación neutra o débil podría sobrevivir en un cromosoma no recombinante como es el cromosoma Y, de otro modo ésta sería eliminada (Charlesworth, 1994; Figura 8).

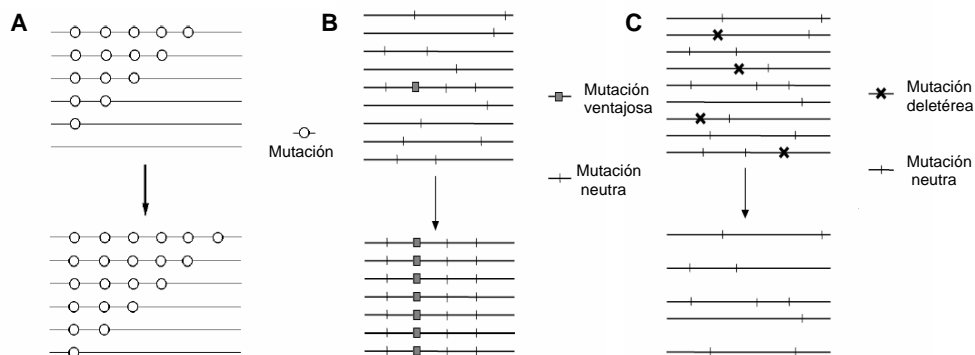
- Inserciones de retrotransposones: los retrotransposones son secuencias de ADN que tienen la propiedad de replicarse y cambiar de posición dentro del genoma. Cuando se retrotransponen pueden arrastrar con ellos secuencias que se incorporaran al nuevo sitio de inserción. Durante la evolución de los homínidos hace 4-5 millones de años, una porción del brazo largo del cromosoma X fue incorporada al cromosoma Y (Page *et al.*, 1984) y fue reposicionada dentro de la zona MSY del cromosoma Y vía recombinación intracromosomal entre retrotransposones LINE (*long-interspersed element*) (Schwartz *et al.*, 1998). La presencia de estos elementos en la zona MSY permite que se produzca la recombinación dentro del cromosoma Y, y que, en consecuencia, ciertas secuencias de ADN cambien su posición. Con este proceso se altera el orden cromosómico, pudiendo generar además la inactivación de genes y generación de pseudogenes.
- Duplicaciones: las duplicaciones surgen cuando un segmento cromosómico se replica más de una vez por error en la duplicación del ADN, como producto de una reorganización cromosómica de tipo estructural o relacionados con un proceso de sobrecruzamiento defectuoso. Las duplicaciones no suelen ser deletéreas; constituyen una fuente de nuevo material genético y base para nuevos cambios evolutivos. Muchas de las familias génicas con un origen evolutivo

común y las familias multigénicas pueden tener su origen en las duplicaciones.

Estos procesos se traducen en una acumulación de mutaciones que inactivan genes, la generación de pseudogenes, un aumento de las secuencias repetitivas y una reducción del tamaño efectivo. Todas estas consecuencias se ven potenciadas por el estado de hemigiosis en el que se encuentra el cromosoma Y.

Se ha especulado con el hecho de que la continua degeneración de la zona MSY podría llevar a la desaparición del cromosoma Y (Graves, 2002). Esta extinción ya ha ocurrido en especies del género *Ellobius* en el que el sexo está controlado por el número de cromosomas X (XX en hembras y XO en los machos) (Graves, 1995; Just *et al.*, 1995). En mamíferos se ha predicho la posibilidad de que se extinga el cromosoma Y, siendo un proceso que ocurriría dentro de unos 10-100 millones de años (Graves, 2002).

**Figura 8.** En A) Ejemplo de Muller Ratchet, en el que se puede apreciar como se acumulan mutaciones, en B) Efecto Hitchhiking donde mutaciones neutras se seleccionan por estar asociadas a una mutación ventajosa y en C) Selección Background en la que mutaciones neutras asociadas a mutaciones deletéreas se van a perder.



### 2.3.3 Estructura del cromosoma Y en mamíferos

El cromosoma Y es un cromosoma único, distinto del resto de cromosomas autosómicos debido a características propias que ningún otro cromosoma presenta, como por ejemplo las largas porciones de ADN especulares o palindrómicas de secuencia casi idéntica y la carencia de un cromosoma homólogo con el que recombinar. Como modelo general de cromosoma Y para explicar su estructura se utilizará el cromosoma Y humano, por ser el más estudiado y el mejor conocido.

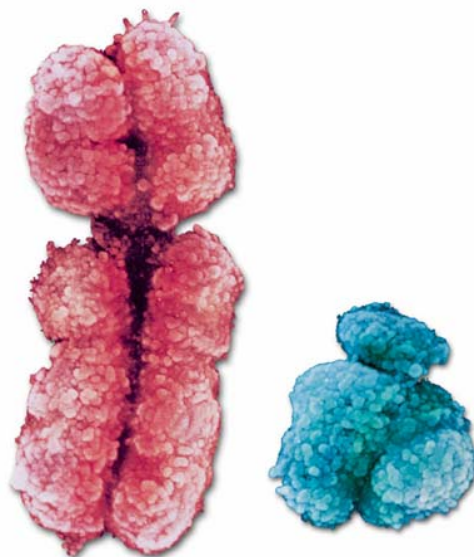
#### 2.3.3.1 El cromosoma Y humano

Los cromosomas sexuales, X e Y, están muy diferenciados (Figura 9). En la especie humana el cromosoma X o HSAX representa el 5% del genoma portando 1100 genes cuyas funciones son una mezcla de funciones de mantenimiento o *housekeeping* y funciones especializadas (Ross *et al.*, 2005). El cromosoma X está altamente conservado, encontrando casi el mismo orden y contenido en genes entre especies (Raudsepp *et al.*, 2004) con ciertas excepciones como es el caso del cromosoma X de ciertos roedores donde los genes no se encuentran en el mismo orden (Bourque *et al.*, 2004). La mayoría de los genes presentes en el cromosoma X no están relacionados con funciones sexuales o reproductivas, aunque estos genes están presentes en mayor frecuencia que en los cromosomas autosómicos (Saifi *et al.*, 1999).

El cromosoma Y humano o HSAY representa el 2% del genoma de la especie con unas 60 Mb. Este cromosoma está compuesto de dos tipos de secuencias: la secuencia eucromática que porta genes activos y las secuencias heterocromáticas que serían ADN no funcional. Estos dos tipos de secuencias se encuentran entremezcladas a lo largo del cromosoma Y. La secuencia eucromática ocupa 23 Mb y el resto consiste en ADN heterocromático repetitivo. La presencia de grandes bloques de secuencias repetitivas, la similitud de estas secuencias y la ausencia de recombinación con el cromosoma X, causó grandes dificultades en el

proyecto del genoma humano (Skaletsky *et al.*, 2003). Este cromosoma necesita ser estudiado de formas diferentes a las del resto del genoma. Las técnicas de secuenciación estándar no funcionan debido a la larga porción de ADN casi idéntico que carece de las señales necesarias para realizar el correcto ensamblaje. En ausencia de recombinación, el mapeo genético no proporciona información y los análisis físicos son de mayor importancia. La disponibilidad de la secuencia casi completa de las porciones eucromáticas del cromosoma podría incluso tener más impacto que la información aportada por la secuencia del resto del genoma de la especie (Jobling y Tyler-Smith 2003). El conocimiento de su secuencia podría ofrecer explicaciones a la distinta susceptibilidad a enfermedades entre hombres y mujeres, explicar ciertas enfermedades ligadas a anomalías en este cromosoma como la azoospermia, la infertilidad (Sun *et al.*, 1999; Kuroda-Kawaguchi *et al.*, 2001) o explicar los fenotipos femeninos (Rao *et al.*, 1997).

**Figura 9.** Cromosomas sexuales humanos (Jegalian y Lahn, 2001, BIOPHOTO ASSOCIATES Photo Researchers, Inc.) en los que se puede apreciar la diferencia de tamaño entre el cromosoma X y el cromosoma Y.





El cromosoma Y contiene una pequeña región que muestra homología con el cromosoma X. Esta región se conoce como zona pseudoautosomal o PAR. La zona PAR representa un papel muy importante durante la meiosis, ya que en ella se produce el sobrecruzamiento y la recombinación con el cromosoma X, resultando esencial para una correcta segregación de los cromosomas sexuales (Ellis y Goodfellow, 1989). La zona PAR se encuentra situada en la región distal del brazo corto de los cromosomas X e Y. En humanos, a diferencia de otros mamíferos, hay además una segunda zona PAR situada en la otra región distal de los cromosomas sexuales. Las zonas pseudoutosomales en humanos se denominan PAR1 y PAR2. El resto del cromosoma Y constituye la zona macho específica (MSY). La zona MSY también se denominó como región no recombinante o NRY, pero este término se ha abandonado al probarse que puede producirse recombinación dentro del cromosoma Y (Skaletsky *et al.*, 2003).

Por lo general, la zona MSY constituye el 95% del cromosoma mientras que el otro 5% se corresponde con la zona PAR. La mayor parte de la zona MSY está compuesta por heterocromatina, ADN repetitivo de secuencia simple que no contiene genes. La eucromatina en el cromosoma Y abarca, aproximadamente, 23 Mb del total de la zona MSY de los que 14,5 Mb están en el brazo largo (Yq) y el resto en el brazo corto (Yp), (Skaletsky *et al.*, 2003). Es una zona de escaso contenido génico, donde pocos genes están activos, de hecho la mayoría de las 156 unidades de transcripción en la zona MSY son pseudogenes. Los 78 genes activos codifican, en conjunto, 27 proteínas distintas. De estas 27 proteínas, 20 tienen su pareja en el cromosoma X (Skaletsky *et al.*, 2003).

### **2.3.3.1.1 Clasificación de las secuencias eucromáticas según su origen**

Las secuencias eucromáticas del cromosoma Y se pueden clasificar en secuencias X- transpuestas, X- degeneradas y amplicónicas. La región eucromática de la zona MSY está compuesta por un mosaico de estas 3 clases de secuencias. El contenido en genes se especifica en la Tabla 2.

Las secuencias X- transpuestas son secuencias que fueron intercambiadas a partir del cromosoma X. Tienen un tamaño total de 3,4 Mb y presentan una identidad del 99% con secuencias de ADN presentes en HSAXq21, una banda situada en la mitad del brazo largo del cromosoma X. Las secuencias X-transpuestas reciben este nombre debido a que su presencia en el MSY humano es el resultado de una masiva transposición del cromosoma X al cromosoma Y hace 3-4 millones de años, después de la divergencia entre humanos y chimpancés (Page *et al.*, 1984; Munn *et al.*, 1997; Schwartz *et al.*, 1998). Se ha observado una inversión que parte el bloque de secuencias X- transpuestas en 2 segmentos no contiguos (Munn *et al.*, 1997; Schwartz *et al.*, 1998). Las secuencias X-transpuestas no participan en el sobrecruzamiento entre el cromosoma X e Y durante la meiosis, lo que las diferencia de las secuencias de la zona PAR. Estas secuencias presentan la densidad más baja de genes dentro de las clases de secuencias presentes en la eucromatina de la región MSY y la densidad más alta de elementos repetitivos intercalados (*interspersed repeat elements*), como el elemento LINE que ocupa el 36% de las secuencias X-transpuestas (Skaletsky *et al.*, 2003; Venter *et al.*, 2001). La secuencia homóloga a las secuencias X-transpuestas, Xq21, también se caracteriza por presentar baja densidad de genes y alta presencia de elementos repetitivos. Dentro de las secuencias X-transpuestas sólo se han identificado 2 genes. Ambos genes presentan homólogos en HSAXq21.

Las secuencias X-degeneradas son reliquias supervivientes de los cromosomas autosómicos de los que derivan los cromosomas sexuales. En contraste con las

regiones X-transpuestas, las secuencias X-degeneradas presentan genes o pseudogenes de copia única homólogos a genes ligados al cromosoma X. Estos genes de copia única y pseudogenes presentan una identidad nucleotídica de entre el 60 y el 96% con sus homólogos en el cromosoma X. Las secuencias X degeneradas codifican 16 de los 27 distintos tipos de proteínas o familias de proteínas de la MSY. Se han encontrado 13 pseudogenes que presentan similitud con exones e intrones del gen funcional del cromosoma X. Otros 14 casos serían genes funcionales que se transcriben y que codifican para proteínas muy similares pero no idénticas isoformas a las codificadas en el cromosoma X. Tanto los pseudogenes como las isoformas no idénticas muestran la degradación constante debido a mutaciones y el proceso de diferenciación surgido entre el cromosoma X y el Y. En total, 12 genes MSY de expresión ubicua están localizados en la región X degenerada. En cambio, de los 11 genes MSY que se expresan predominantemente en testículo sólo uno, el gen de determinación del sexo (SRY), es X- degenerado (Skaletsky *et al.*, 2003).

La tercera clase de secuencias eucromáticas son las amplicónicas. Estas secuencias se encuentran dentro de segmentos palindrómicos múltiples y repetitivos. Están compuestas en gran parte por secuencias que muestran una marcada similitud (un 99%) con otras secuencias de miles e incluso cientos de kb de la zona MSY. Son unidades MSY específicas de repetición de las que hay muchas familias llamadas amplicones. Los amplicones están localizados en 7 segmentos diseminados a lo largo de un total de 10,2 Mb de la eucromatina del brazo largo y la zona proximal del brazo corto. Las regiones amplicónicas presentan la mayor densidad de genes, codificantes y no codificantes, respecto a las tres clases de eucromatina MSY. En esta región se han identificado 9 familias distintas de genes MSY específicas que codifican proteínas. Estas familias de genes presentan un número de copias que va desde las 2 copias en los genes VCY, XKRY, HSFY y PRY, 3 copias el gen BPY2, 4 copias en los genes CDY, y DAZ,

6 copias del gen RBMY o aproximadamente 35 copias del gen TSPY. Estas 9 familias abarcan aproximadamente 60 unidades de transcripción. Además, las secuencias amplicónicas incluyen al menos otras 75 unidades de transcripción que no codifican proteína (Skaletsky *et al.*, 2003). De esas 75 unidades de transcripción no codificantes, 65 son miembros de 15 familias específicas MSY, y las restantes están presentes como copia única. Considerando conjuntamente elementos codificantes y no codificantes, las secuencias amplicónicas contiene 135 de las 156 unidades de transcripción descritas. Los genes situados en secuencias amplicónicas y las unidades de transcripción muestran una expresión limitada. Las 9 familias de proteínas se expresan predominante o exclusivamente en testículo así como la mayoría de las unidades de transcripción no codificantes (Skaletsky *et al.*, 2003). De las 3 clases de secuencias eucromáticas, las amplicónicas presentan las mas bajas densidades de elementos LINE1 y elementos repetitivos dispersos. De hecho, el contenido de estos elementos repetitivos dispersos en la zona MSY amplicónica es menor que la contenida en el total del genoma humano (un 35% frente a un 44%; Skaletsky *et al.*, 2003). La existencia de zonas palindrómicas posibilita que genes del cromosoma Y situados en esas zonas puedan recombinar con genes del cromosoma Y situados en el palíndromo idéntico por lo que existe recombinación dentro del cromosoma Y.

Las secuencias X-transpuestas y X-degeneradas reflejan el origen común y antiguo de los dos cromosomas y muestran la evidencia de una evolución por etapas a través del tiempo.

**Tabla 2.** Genes presentes en el cromosoma Y humano clasificados según su origen (Skaletsky et al., 2003)

Secuencia MSY	Símbolo	Nombre del gen	Nº de copias	Tejido de expresión	Homólogo en el X	Homólogo autosomal
X-transpuesta	TGIF2LY	<i>TGF (beta)-induced transcription factor 2-like Y</i>	1	Testículo	TGIF2LX	–
	PCDH11Y	<i>Protocadherin 11 Y</i>	1	Cerebro fetal y cerebro	PCDH11X	–
Total			2			
X-degenerada	SRY	<i>Sex determining region Y</i>	1	Testículo predominantemente	SOX3	–
	RPS4Y1	<i>Ribosomal protein S4 Y isoform 1</i>	1	Ubicuo	RPS4X	–
	ZFY	<i>Zinc finger Y</i>	1	Ubicuo	ZFX	–
	AMELY	<i>Amelogenin Y</i>	1	Dientes	AMELX	–
	TBL1Y	<i>Transducin (beta)-like 1 protein Y</i>	1	Cerebro fetal y próstata	TBL1X	–
	PRKY	<i>Protein kinase Y</i>	1	Ubicuo	PRKX	–
	USP9Y	<i>Ubiquitin-specific protease 9 Y</i>	1	Ubicuo	USP9X	–
	DBY	<i>Dead box Y</i>	1	Ubicuo	DBX	–
	UTY	<i>Ubiquitous TPR motif Y</i>	1	Ubicuo	UTX	–
	TMSB4Y	<i>Thymosin (beta)-4 Y</i>	1	Ubicuo	TMSB4X	–
	NLGN4Y	<i>Neuroigin 4 isoform Y</i>	1	Cerebro fetal, cerebro, próstata y testículo	NLGN4X	–
	CYorf15A	<i>Chromosome Y open reading frame 15A</i>	1	Ubicuo	CXorf15	–
	CYorf15B	<i>Chromosome Y open reading frame 15A</i>	1	Ubicuo	CXorf15	–
	SMCY	<i>SMC (mouse) homologue, Y</i>	1	Ubicuo	SMCX	–
	EIF1AY	<i>Translation initiation factor 1A Y</i>	1	Ubicuo	EIF1AX	–
	RPS4Y2	<i>Ribosomal protein S4 Y isoform 2</i>	1	Ubicuo	RPS4X	–
Total			16		–	
Ampliconica	TSPY	<i>Testis-specific protein Y</i>	~35	Testículo	VCX	–
	VCY	<i>Variable charge Y</i>	2	Testículo	–	–
	XKRY	<i>XK related Y</i>	2	Testículo	–	–
	CDY	<i>Chromodomain Y</i>	4	Testículo	–	CDYL
	HSFY	<i>Heat shock transcription factor Y</i>	2	Testículo	–	–
	RBMV	<i>RNA-binding motif Y</i>	6	Testículo	RBMX	–
	PRY	<i>PTP-BL related Y</i>	2	Testículo	–	–
	BPY2	<i>Basic protein Y 2</i>	3	Testículo	–	–
	DAZ	<i>Deleted in azoospermia</i>	4	Testículo	–	DAZL
Total			~60			
Total global			~78			

#### **2.3.3.1.2 Clasificación de las secuencias eucromáticas según sus características**

Además de por su origen, los genes presentes en el cromosoma Y humano se pueden clasificar teniendo en cuenta las características del gen. En base a este criterio dividimos el contenido génico del cromosoma Y en cuatro grupos:

Grupo I: incluye genes cuya función está relacionada con la determinación sexual (SRY y ZFY).

Grupo II: formado por 16 genes (EIF1AY, CYorf15A y 15B, DBY, NLGN4Y, PCDH11Y, PRKY, USP9Y, RPS4Y1, RPS4Y2, SMCY, TBL1Y, TGIF2LY, TMSB4Y, CSF2RA, ANT3 y UTY) de expresión ubicua, la mayoría de ellos relacionados con funciones de mantenimiento o *housekeeping*.

Grupo III: contiene 9 genes (RBMX, DAZ, TSPY, CDY, BPY2, XKRY, PRY, HSFY y VCY) cuya característica es estar presentes en múltiples copias en la zona MSY. Estos genes se expresan sólo en el testículo y están relacionados con la espermatogenesis y la fertilidad (Lahn y Page, 1997).

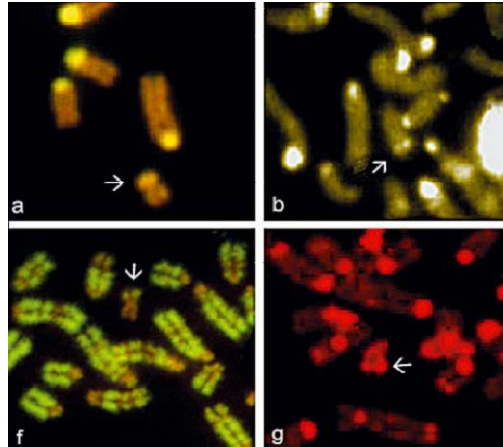
Grupo IV: incluye los genes AMELY y GCY relacionados con el control del crecimiento embrionario, estatura y desarrollo dental (Fincham *et al.*, 1990; Kirsch *et al.*, 2000, 2004; Ogata y Matsuo, 1993).

#### **2.3.3.2 El cromosoma Y bovino.**

Al igual que sucede en humanos los cromosomas X e Y están muy diferenciados en Bovino (Figura 10). Del cromosoma X se conoce su contenido génico y su secuencia. El cromosoma X bovino o BTAX es metacéntrico mide 146,5 Mb, se ha mapeado (ej. Ihara *et al.*, 2004), y su secuencia ha sido desvelada recientemente (Bovine HapMap Consortium, 2009; Tellam *et al.*, 2009). El BTAX

presenta completa homología con el HSAX, aunque se observaron reordenamientos internos (Wind *et al.*, 2005). Respecto del BTAY no se ha secuenciado por lo que hay un gran desconocimiento en cuanto a contenido génico y estructura.

**Figura 10.** Cromosomas Y con distintas tinciones de *B. taurus* (a y f) y *B. indicus* (b y g) señalados con una flecha (Bar=10)  $\mu$ m. (Di Meo *et al.*, 2005).



El cromosoma Y de *B. taurus* y de otras especies bovinas como el gaur (*B. gaurus*) y el banteng (*B. javanicus*) es metacéntrico, a diferencia de los autosomas de estas especies que son acrocéntricos. El cromosoma Y de la otra especie de bovino doméstico, el cebú (*B. indicus*), es acrocéntrico (Goldammer *et al.*, 1997). Esta última es la morfología más probable para el cromosoma Y ancestral de los Bovini (Gallagher *et al.*, 1999). Entre los cromosomas Y de *B. taurus* y *B. indicus* ha sido identificada una transposición centromérica o una inversión pericéntrica (Di Meo *et al.*, 2005). El cromosoma Y bovino representa tan solo el 1,7 % del genoma de la especie. A diferencia del cromosoma Y humano presenta una sola región PAR (Skaletsky *et al.*, 2003; Liu *et al.*, 2007a), situada en la región distal del brazo corto del cromosoma (Yp), ocupando el resto del cromosoma la zona MSY. El tamaño estimado para el cromosoma Y bovino es de 51 Mb y la región eucromática representa el 50 % del cromosoma (Liu *et al.*, 2007a). Dentro de la región MSY, la parte localizada en el brazo corto (Yp) y la región proximal del

brazo largo (Yq) son regiones eucromáticas. El resto del brazo Yq contiene regiones heterocromáticas, ricas en secuencias repetitivas.

Debido al desconocimiento de la secuencia de este cromosoma, el proceso de localización de genes en el cromosoma Y bovino todavía se encuentra en desarrollo. El gen SRY (Daneau *et al.*, 1995) se localiza en la región Yq (Iannuzzi *et al.*, 2001b). Otros genes X degenerados de copia única, como el AMELY (amelogenina Y) (Gibson *et al.*, 1991) y ZFY (*zinc finger protein, Y-encoded*) (Xiao *et al.*, 1998), también han sido caracterizados. El gen TSPY (Jakubiczka *et al.*, 1993; Vogel *et al.*, 1997 a, b) está situado en la región amplicónica presentando un número de copias que oscila entre las 50 y las 200 repeticiones. Verkaar *et al.*, (2004) estimaron el número de copias de este gen en 90 loci, la mayoría de ellos situados en el brazo Yp. La mitad de estos *loci* codificarían para una proteína funcional mientras que la otra mitad serían pseudogenes.

La localización de genes en el cromosoma Y bovino se está llevando a cabo mediante técnicas de genómica comparativa, a partir de las secuencias de genes presentes en el cromosoma Y de otras especies, fundamentalmente la especie humana. Sin embargo esta estrategia no parece muy útil ya que no siempre se obtienen los resultados esperados. Es el caso del gen DAZL (*Deleted in Azoospermia Like*) que en la especie humana presenta 4 copias en el cromosoma Y, y una copia adicional en el cromosoma 3 (Saxena *et al.*, 2000). Sin embargo, en bovinos el gen DAZL está presente en el cromosoma 1 y en forma de pseudogen en el cromosoma 16, pero no está presente en el cromosoma Y (Liu *et al.*, 2007). Algo similar ocurrió respecto del gen CDYL (*Chromodomain-Y-like*), que, aunque en la especie humana se encuentra en el cromosoma Y además de en dos autosomas (Lahn y Page, 1999), en bovino se localiza en los cromosomas 24 y 18 pero no en el cromosoma Y (Wang *et al.*, 2008).



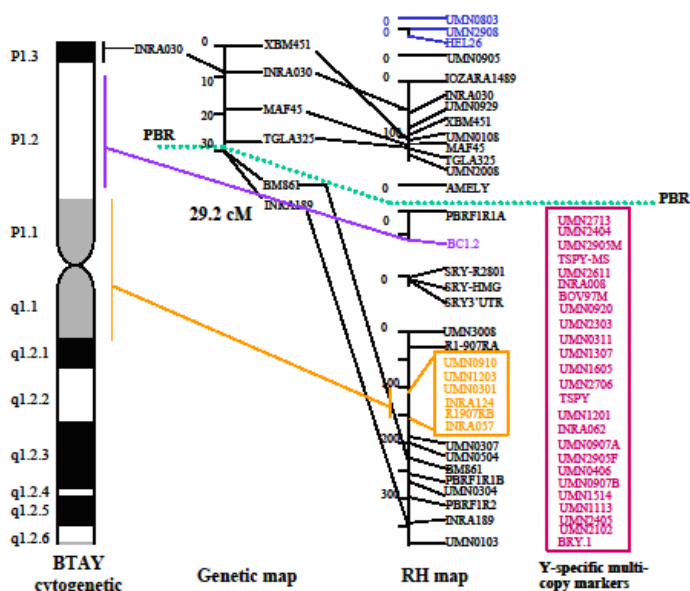
Al igual que acontece en la especie humana, anormalidades en el cromosoma Y están asociadas a fenotipos femeninos (Iannuzzi *et al.*, 2001a) y azoospermia (Iannuzzi *et al.*, 2001b).

#### **2.3.3.2.1 Mapa del cromosoma Y bovino**

El mapa del cromosoma Y bovino es el más avanzado entre las especies de ganado con más de 100 marcadores identificados incluyendo en esta lista genes, SNP y STR (Liu *et al.*, 2007).

Numerosos microsatélites han sido mapeados por análisis de ligamiento en la región PAR bovina (Sonstegard *et al.*, 2001). Sin embargo, la ausencia de recombinación meiótica en la zona MSY, hace del mapeo por híbridos de radiación (RH) la técnica de elección para caracterizar la estructura de esta región. Existen en la actualidad tres mapas RH para el BTAY (Liu *et al.*, 2002; Sonstegard *et al.*, 2001; Ihara *et al.*, 2004) (ver Figura 11). Estos mapas son útiles para el clonado posicional de genes Y específicos e identificar posibles QTL del cromosoma Y (Liu *et al.*, 2002). Varios microsatélites MSY específicos no *multilocus* identificados por Liu *et al.*, (2002), han sido localizados por mapeo de hibridación radiactivo (*radiation hybrid mapping*; Liu *et al.*, 2003). El microsatélite IDGVA50 ha sido localizado por hibridación de fluorescencia in situ (FISH *fluorescent in situ hybridization*) en la región Yq (Iannuzzi *et al.*, 2001b) asociado al *loci* TSPY (Verkaar *et al.*, 2004).

**Figura 11.** Mapa RH del cromosoma Y bovino (Liu *et al.*, 2003). En la izquierda se representa un ideograma del bandeo G+C del cromosoma Y, en el medio se representa un mapa de ligamiento de la zona PAR. La línea de puntos con la leyenda PBR indica la supuesta ubicación la PAB (límites entre la zona MSY y PAR). Los marcadores en el recuadro amarillo del mapa RH son centroméricos. Los marcadores del recuadro rosa son marcadores multilocus Y específicos, no incluidos en el mapa RH.



Los límites entre la zona MSY y PAR se denominan PAB o *pseudoautosomal boundary*. Esta zona se ha caracterizado recientemente en *B. taurus* y otras especies. A partir del gen Amelogenina Y, considerado el gen Y específico localizado mas próximo a la PAB (Lui *et al.*, 2002), se han determinado la zona PAB (Van Laere *et al.*, 2009). En el mismo proceso, se han caracterizado 4 genes de la zona MSY que se hallarían próximos a la zona PAB. La secuencia de estos genes sería USP9Y- OFD1Y-AMELY-EIF1AY- zona PAB. Tanto en el cromosoma Y como en el X, la región PAB esta flanqueada por segmentos de 1 a 5kb que presentan un alto contenido en G+C. En el cromosoma Y, está además flanqueada por un segmento de unas 17 kb compuesto por elementos MaLRs

(*Mammalian apparent LTR-retrotransposons*) que son un tipo de retrotransposones que presentan largas repeticiones terminales o LTR) (Van Laere *et al.*, 2009).

#### **2.3.3.2.2 Microsatélites descritos en el cromosoma Y bovino**

Los marcadores del cromosoma Y por el momento son escasos y la mayoría de ellos se aislaron en la región PAR, son monomórficos o presentan un elevado número de copias (Liu *et al.*, 2002; Van Hooft *et al.*, 2002). En la Tabla 3 se detallan todos los microsatélites descritos hasta el momento en el cromosoma Y bovino. Los marcadores que presentan un elevado número de copias se denominan microsatélites *multilocus* o *interspersed multilocus microsatellites* (IMM). Una de las características más representativas del cromosoma Y es su naturaleza repetitiva, lo que hace que el descubrimiento de marcadores útiles en especies no modelo sea muy difícil. La naturaleza repetitiva del cromosoma Y se postula como un mecanismo para mantener los genes específicos del cromosoma Y en un ambiente no recombinante que podría erosionar al cromosoma Y con un número elevado de mecanismos entre los que se incluye el estado de hemizigosis (Vallender y Lahn 2004; Murphy *et al.*, 2006).

**Tabla 3.** *Microsatélites descritos en el cromosoma Y bovino. Para algunos de los marcadores se especifica su posición en el cromosoma Y, siendo Yp el brazo corto e Yq el brazo largo del cromosoma Y.*

Nombre del Locus	Mapa	Marcador	Referencia
DYS23	Y	microsatelite (IOBT1489)	Barendse et al. 1997
DYS3	Y	microsatelite (INRA008)	Vaiman et al. 1994
DYS4	Y	microsatelite (INRA057)	Vaiman et al. 1994
DYS5	Y	microsatelite (INRA062)	Vaiman et al. 1994
DYS6	Y	microsatelite (INRA124)	Vaiman et al. 1994
DYS7	Y	microsatelite (INRA126)	Vaiman et al. 1994
DYS8	Y	microsatelite (BM861)	Kappes et al. 1997
DYZ10	Yq	microsatelite (IDVGA50)	Mezzelani et al. 1995
BL22A	Yp	microsatelite	Barendse et al. 1997; Liu et al. 2002
BL22B	Yp	microsatelite	Barendse et al. 1997; Liu et al. 2002
BYM-1	Y	microsatelite BYM-1	Matthews et al. 1991; Ward et al. 2001
HEL26 (DXS29)	Yp	microsatelite	Vilki et al. 1995
INRA189	Yq	microsatelite (INRA189)	Kappes et al. 1997
INRA30	Yp13	microsatelite (INRA030)	Kappes et al. 1997
MAF45	Yp	microsatelite (MAF45,ovine)	Kappes et al. 1997
MCM74	Y	microsatelite (MCM74)	Sonstegard et al. 2001
TGLA325	Yp	microsatelite (TGLA325)	Kappes et al. 1997
UMN0103	Y	microsatelite	Liu et al. 2002
UMN0108	Yp	microsatelite	Liu et al. 2002
UMN0301	Y	microsatelite	Liu et al. 2002
UMN0304	Y	microsatelite	Liu et al. 2002
UMN0307	Y	microsatelite	Liu et al. 2002
UMN0311	Y	microsatelite	Liu et al. 2002
UMN0406	Y	microsatelite	Liu et al. 2002
UMN0504	Y	microsatelite	Liu et al. 2002
UMN0705	Y	TSPY-microsatelite	Liu et al. 2002
UMN0803	Yp	microsatelite	Liu et al. 2002
UMN0905	Yp	microsatelite	Liu et al. 2002
UMN0907A	Y	microsatelite	Liu et al. 2002
UMN0907B	Y	microsatelite	Liu et al. 2002
UMN0910	Y	microsatelite	Liu et al. 2002
UMN0920	Y	microsatelite	Liu et al. 2002
UMN0929	Yp	microsatelite	Liu et al. 2002
UMN1113	Y	microsatelite	Liu et al. 2002
UMN1201	Y	microsatelite	Liu et al. 2002
UMN1203	Y	microsatelite	Liu et al. 2002
UMN1307	Y	microsatelite	Liu et al. 2002
UMN1514	Y	microsatelite	Liu et al. 2002
UMN1605	Y	microsatelite	Liu et al. 2002
UMN2008	Yp	microsatelite	Liu et al. 2002
UMN2102 (BTGA50)	Y	microsatelite	Liu et al. 2002
UMN2303	Y	microsatelite	Liu et al. 2002
UMN2404	Y	microsatelite	Liu et al. 2002
UMN2405 (BTMS2437)	Y	microsatelite	Liu et al. 2002
UMN2611	Y	microsatelite	Liu et al. 2002
UMN2706	Y	microsatelite	Liu et al. 2002
UMN2713	Y	microsatelite	Liu et al. 2002
UMN2905F	Y	microsatelite	Liu et al. 2002
UMN2905M	Y	microsatelite	Liu et al. 2002
UMN2908	Yp	microsatelite	Liu et al. 2002
UMN3008	Y	microsatelite	Liu et al. 2001
XBM31	Yp	microsatelite (XBM31)	Ponce de León et al. 1996
XBM451	Yp	microsatelite (XBM451)	Kappes et al. 1997

### 2.3.3.2.3 Análisis de la variabilidad genética del cromosoma Y bovino

Si hay ciertas carencias de datos en cuanto a la información que nos ofrecen las líneas maternas en ganado bovino, la información obtenida hasta el momento para las líneas paternas es todavía menor.

Los primeros estudios filogenéticos en cromosoma Y de ganado se hicieron con un pequeño número de microsatélites. Se conocen pocos marcadores microsatélites (Bishop *et al.*, 1994; Vaiman *et al.*, 1994; Mezzelani *et al.*, 1995; Kappes *et al.*, 1996; Liu *et al.*, 2003) y sólo 5 SNP (Götherström *et al.*, 2005) específicos del cromosoma Y han sido aislados. Los marcadores INRA124 y INRA126 desarrollados por Vaiman *et al.*, (1994) son dos de los más utilizados para hacer estudios de introgresiones de *B. indicus* en *B. taurus*. El marcador polimórfico INRA124 fue utilizado por Hanotte *et al.*, (2000), empleado en los trabajos de Edwards *et al.*, 2000 y 2007 y también usado por Giovambattista en 2000. En 2007, Li *et al.*, utilizan los marcadores INRA124, INRA126, INRA189, BM861 y BYM-1 para diferenciar *B. indicus* y *B. taurus*. Todos estos estudios permiten diferenciar a *B. indicus* y *B. taurus* gracias al marcador INRA 124. Este marcador presentaba dos tamaños, 132 para *B. indicus* y 130 en *B. taurus*. Otros microsatélites utilizados para evaluar la variabilidad genética del cromosoma Y del ganado doméstico son el UMN0103 y el UMN2404 (Lui *et al.*, 2003), ambos marcadores presentan más de una copia (Cai *et al.*, 2006). El UMN0103 presenta unas pocas bandas en el rango de 124 a 136 pb. El UMN2404 presenta un patrón de bandas en un rango se las 85 pb a las 112 pb. Estos dos microsatélites permiten diferenciar también *B. indicus* de *B. taurus*. Todos estos estudios han sido realizados en áreas restringidas, no llegándose a hacer un análisis geográfico global.

En 2005 Götherström *et al.*, encuentran 5 SNP (DBY1, DBY7, UTY19, ZFY4 y ZFY5), un microsatélite (DBY1) y una delección (ZFY5-indel) que permiten

distinguir 3 haplogrupos en el cromosoma Y bovino, el Y1, el Y2 y el Y3. Los dos primeros haplogrupos están presentes en *B. taurus*, mientras que el tercero, el Y3, está presente solamente en *B. indicus*. El haplogrupo Y1 se correspondería con muestras procedentes del norte de Europa, mientras que el haplogrupo Y2 se correspondería con muestras procedentes del sur de Europa. Götherström *et al.*, (2005) utilizaron en su trabajo ADN antiguo, sugiriendo que los individuos del haplotipo Y2 se corresponderían con los individuos domesticados en el cercano oriente mientras que los individuos del haplotipo Y1 serían resultado de la introgresión de uros locales en el ganado domesticado. Sin embargo, Bollongino *et al.*, (2008) usando el SNP UTY 19 y la delección ZFY5-indel y Svensson y Götherström (2008) usando el SNP UTY19, realizan estudios con ADN antiguo de uros medievales procedentes de distintas partes de Europa comparándolas con muestras modernas. Los resultados de ambos trabajos no apoyan la hipótesis presentada por Götherström *et al.*, (2005) respecto a los orígenes de los haplogrupos Y1 e Y2.

Recientemente, Ginja *et al.*, (2008) realizan un estudio combinando información proveniente de los SNP descritos por Götherström *et al.*, (2005) y un grupo de 7 microsatélites (DBY1, BM861, INRA124, INRA189, UMN0103, UMN0307 y UMN0504) de los cuales 6 son polimórficos (todos a excepción del UMN0504). Este estudio desde el punto de vista metodológico es el que aportaría más información, ya que combina información obtenida con SNP y microsatélites. Desde el punto de vista geográfico se trata de un trabajo realizado en un área geográfica restringida, Portugal, por lo que quedan muchas incógnitas que despejar.

La parte principal de esta Memoria es el siguiente compendio de de publicaciones científicas realizadas y que figuran como aval del trabajo realizado para la consecución de los objetivos propuestos:

Pérez-Pardal, L., Royo, L.J., Álvarez, I., Ponce de León, F.A., Fernández, I., Casais, R., Goyache, F. (2009) Female segregation patterns of the putative Y-chromosome specific microsatellite markers INRA124 and INRA126 do not support their use for cattle population studies. *Animal Genetics*, 40: 560-564. doi:10.1111/j.1365-2052.2009.01870.x

Pérez-Pardal, L., Royo, L.J., Beja-Pereira, A., Curik, I., Traoré, A., Fernández, I., Sölkner, J., Álvarez, I., Bozzi, R., Chen, S., Ponce de León, F.A., Goyache, F. (2010) Y-specific microsatellites reveal an African subfamily in taurine (*B. taurus*) cattle. *Animal Genetics*, 41:232-241.

Pérez-Pardal, L., Royo, L.J., Beja-Pereira, A., Chen, S., Cantet, R.J.C., Traoré, A., Curik, I., Sölkner, J., Bozzi, R., Fernández, I., Álvarez, I., Gutiérrez, J.P., Gómez, E., Ponce de Leon, F.A., Goyache, F. (2010) Multiple paternal origins of domestic cattle revealed by Y-specific interspersed multilocus microsatellites. *Heredity*, aceptado para publicación. doi:10.1038/hdy.2010.30

Pérez-Pardal, L., Ginja, C., Royo, L.J., Fernández, I., Álvarez, I., Beja-Pereira, A., Traoré, A., Cantet, R.J.C., Ponce de Leon, F.A., Penedo, C., Goyache, F. (2010) Genetic structure of the bovine Y- specific microsatellite UMN0103 reflects the genetic history of the species. Enviado para su publicación a la revista *Animal Genetics*.

Por otro lado, estos trabajos se han llevado a cabo en conexión con otros que no están incluidos en la presente Memoria y entre los que destacan:

Pajares, G., Balseiro, A., **Pérez-Pardal, L.**, Gamarra, J.A., Monteagudo, L., Goyache, F., Royo, L.J. (2009) Sry-negative XX true hermaphroditism in a roe deer: a case report. *Animal Reproduction Science*, 112, 190-197. doi:10.1016/j.anireprosci.2008.04.018

**Pérez-Pardal, L.**, Royo, L.J., Álvarez, I., Fernández, I., Gutiérrez, J.P., Goyache F. (2007) Caracterización preliminar de las líneas paternas de la raza ovina Xalda de Asturias mediante polimorfismos en el cromosoma Y. *ITEA*, 28, Vol. Extra II, 378-380.





*Artículos*





## Female segregation patterns of the putative Y-chromosome-specific microsatellite markers *INRA124* and *INRA126* do not support their use for cattle population studies

L. Pérez-Pardal\*, L. J. Royo\*, I. Álvarez\*, F. A. Ponce de León†, I. Fernández\*, R. Casais† and F. Goyache\*

\*SERIDA-Somío, C/ Camino de los Claveles 604, E-33203 Gijón (Asturias), Spain. †Department of Animal Science, University of Minnesota, 480 Haecker Hall, 1364 Eckles Avenue, St Paul, MN 55108, USA. ‡SERIDA-Jove, Jove del Medio, E-33299, Gijón (Asturias), Spain

### Summary

Here we tested the segregation and paternal compatibility of markers *INRA124* and *INRA126* on female DNA in 10 different cattle families, in order to clarify the usefulness of these microsatellites for the study of male-mediated population processes in cattle. Their performance was compared with that of four microsatellites located in the PAR-BTAY (*UMN0108*, *UMN0803*, *UMN0929* and *UMN0905*) and another one male-specific microsatellite (*INRA189*). *INRA124* and *INRA126* amplified the same sized fragment in both sexes. Same size alleles were sequenced and the high homology found allowed us to rule out non-specific female amplification. *INRA124* showed full parental compatibility, whilst the locus *INRA126* showed 55% parental incompatibility. Based on these observations, it is recommended that markers *INRA124* and *INRA126* should not be used in studies to characterize male-mediated genetic events in cattle.

**Keywords** cattle, *INRA124*, *INRA126*, phylogeny, Y-chromosome, Y-specific markers.

Assessment of neutral genetic variability on the Y-chromosome in livestock is of high interest to ascertain male genetic legacies throughout the history of a species (MacHugh & Bradley 2001). The number of studies on bovine Y-chromosome (BTAY) polymorphisms has quickly increased (Vaiman *et al.* 1994; Hanotte *et al.* 1997; Liu *et al.* 2003). Such studies have focused on the assessment of male-mediated migration patterns and admixture between *Bos taurus* and *Bos indicus* (Hanotte *et al.* 2000; Li *et al.* 2007). The number of polymorphic-specific markers identified on BTAY is limited (Vaiman *et al.* 1994; Kappes *et al.* 1997; Liu *et al.* 2003) and most population genetic studies in cattle have included *INRA124* and *INRA126* (Edwards *et al.* 2000, 2007; Giovambattista *et al.* 2000; Hanotte *et al.* 2000; Li *et al.* 2007), which were identified and reported to be male-specific by Vaiman *et al.* (1994). Indeed, the chromosomal location of the X- and Y-specific regions on both BTAX and BTAY has remained controversial (Liu *et al.* 2003).

This research aims to clarify the usefulness of the markers *INRA124* and *INRA126* for the study of male-mediated

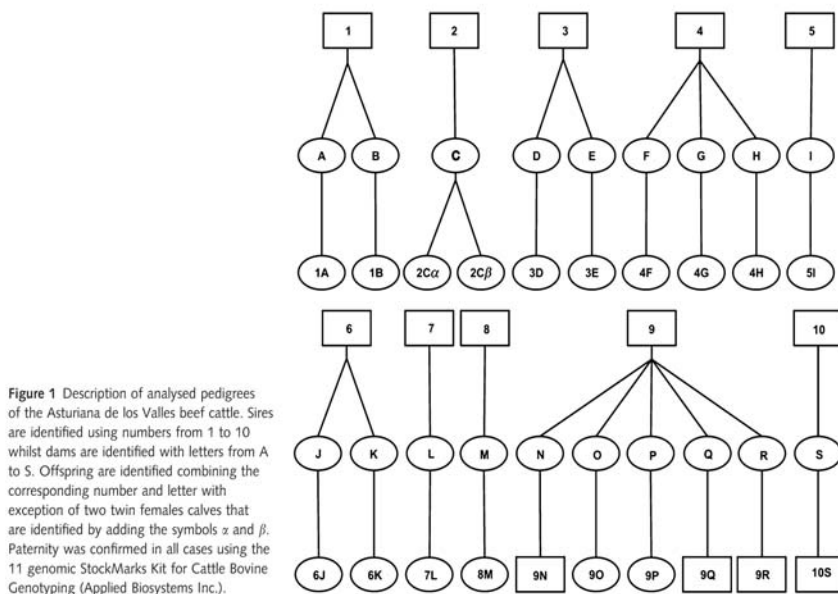
population processes in cattle. This is justified by the uncertainty of their physical location on BTAY and the lack of systematic tests for the amplification of these markers on female DNA. Here, we test the amplification of these markers on female DNA and their segregation and paternal compatibility (markers clonally inherited from father to son) on 10 different cattle families (see Pedigrees in Fig. 1). Their performance is compared with that of five BTAY microsatellites, four located in the pseudo-autosomal region (PAR) of BTAY (*UMN0108*, *UMN0803*, *UMN0929* and *UMN0905*; Liu & Ponce de León 2007) and another male-specific microsatellite (*INRA189*; Kappes *et al.* 1997) (Table 1). This will contribute to the ascertainment of the ability of these markers to reflect only the evolutionary histories of male lineages.

As expected, *INRA189* amplified only from males (Table 2). The four microsatellites located in the PAR-BTAY (*UMN0108*, *UMN0803*, *UMN0905* and *UMN0929*) amplified alleles of the same size range both on male and female DNA, showing full parental compatibility except for *UMN0108* and *UMN0905* (see Table 2, Appendix S1). *INRA124* amplified in both sexes with alleles of the same sizes and showed full parental compatibility. *INRA126* showed a more erratic behaviour. It amplified in both sexes with alleles of the same sizes, but the frequency of parental incompatibility was 55% (see Table 2). Alleles of the same

Address for correspondence

F. Goyache, SERIDA-Somío, C/ Camino de los Claveles 604, E-33203 Gijón (Asturias), Spain.  
E-mail: fgoyache@serida.org

Accepted for publication 18 January 2009

**Table 1** Description of the microsatellite markers tested.

Marker	Primer sequences [5'-3']	No. alleles <sup>1</sup>	Size range <sup>1</sup>	Location <sup>2</sup>	Reference
UMN0108	GATCCATCCACATTGCTCCA-up CCAAGCGTCATCAATTAC-dn	5	190-208	PAR	Liu <i>et al.</i> (2003)
UMN0803	GATCCACATCCCCCTCAC-up CTGCTTCTCTGTCCGCTAA-dn	2	260-279	PAR	
UMN0929	ACCAGCTGATACACAAGTGC-up GGTCAGAGAATGAAACAGAG-dn	8	174-196	PAR	
UMN0905	ATCAACCGTGGTAGCTCTAA-up CTAGAATGTAAACAGCTGC-dn	4	158-164	PAR	
INRA124	GATCTTTGCAACTGGTTTG-up AGGACACAGGTCTGAGAATG-dn	2	130-132	Y	Vaiman <i>et al.</i> (1994)
INRA126	GTTGTGCTCTGCAGAGTAGG-up GACACTCTTTCTATTTCAAGG-dn	5	179-187	Y	Vaiman <i>et al.</i> (1994)
INRA189	TTTGTGTTCCCGTCTGAG-up GAACCTCGTCTCTTGATGCC-dn	2	152-160	MSY	Kappes <i>et al.</i> (1997)

<sup>1</sup>No. alleles and size range found in the 10 cattle families tested in this study.<sup>2</sup>Following Liu & Ponce de León (2007): PAR and MSY mean the pseudoautosomal and male-specific regions of BTAY respectively.

size, on loci *INRA124* and *INRA126*, were sequenced from two individuals of each sex (including a female zebu in *INRA124*). For *INRA124*, we found one polymorphism with respect to the published sequence (Genbank accession no. X71546) (Appendix S1). For *INRA126*, three polymorphisms were found with respect to the published sequence (Genbank accession no. X71553) (Appendix S1). Altogether these results allowed us to rule out non-specific female amplification.

Liu *et al.* (2003) suggested three explanations as to why a Y-specific marker amplifies the same product from both male and female DNA: (i) the common origin of sexual chromosomes, with X- and Y-specific regions still sharing homologous sequences across the genome; (ii) the existence of multiple copies of a given BTAY-specific marker along the chromosome; and (iii) non-specific PCR amplification of female DNA samples. If amplification products of BTAY microsatellites on female DNA differ in size to those obtained

**Table 2** Identification of genotyped individuals and their genotypes for the seven Y-chromosomes tested. The male calf offspring is that amplifying for the microsatellite INRA189. Genotypes showing locus incompatibility are in bold. Each sample was genotyped three times for all markers using different PCR amplifications to verify consistency of obtained results (Appendix S1). Identical results were found regardless of PCR condition.

Identification			INRA189			UMN07108			UMN0803			UMN0905			UMN0929			INRA124			INRA126		
Bull	Cow	Offs	B	C	O	B	C	O	B	C	O	B	C	O	B	C	O	B	C	O	B	C	O
1	A	1A	152	--	190-194	190-194	190-194	279	279	279	164	164-168	164	164-168	174-192	174-176	174-192	130-132	130	181	187	181	
	B	1B	--	--	190-194	200-210	194-210	279	279	279	164	164	164	164	174-192	176-192	174-192	130-132	130	181	181	181	
2	C	2C $\alpha$	152	--	194	190-198	279	279	279	279	164	166	164	164	174-192	176-184	174-176	130-132	130	187	181	187	
		2C $\beta$	--	--	--	190-194	194-198	279	279	279	164	166	164	164	174-192	176-184	174-176	130-132	130	187	181	187	
3	D	3D	152	--	194-202	194-208	194-202	279	279	279	164-166	166	164-166	166	174-192	176-192	174-192	130-132	130	181	185	183	
	E	3E	--	--	--	190-194	194-202	279	279	279	164-166	166	164-166	166	174-192	176-194	176-194	130-132	130	181	185	183	
4	F	4F	152	--	190-194	192-202	192-194	279	279	279	164	164	164	164	174-192	180-186	174-186	130-132	132	183	183	183	
	G	4G	--	--	194-208	194	194-208	279	279	279	166	164-166	164	166	188-194	192-194	188-194	130-132	132	183	183	183	
	H	4H	--	--	--	194-202	190-202	279	279	279	166	164	164	166	176-192	192	176-192	132	130-132	183	181	181	
5	I	5I	152	--	194-200	200-208	200	260-279	260-279	260-279	164	166	166	166	174-194	176-192	192-194	130-132	130	185	179	185	
	J	6J	152	--	192-200	194	192-194	279	260-279	279	164	164	164-166	164	174-194	174-176	174-194	130-132	130	187	181	181	
	K	6K	--	--	192-200	192-194	192	279	279	279	164-166	164	164-166	164	174-180	180-194	174-180	130	130	183	183	183	
7	L	7L	152	--	190-194	192-194	190	260-279	279	260-279	166	164-166	166	164-166	174-192	174	174	130-132	130	183	183	183	
8	M	8M	152	--	--	192-194	194	192-194	279	279	166	164	164	164-166	174-184	174	174	130-132	130-132	130	181	181	
9	N	9N	160	--	160	190-194	194-200	190-194	279	279	158-166	164-166	158-166	176-192	188-192	192	192	130-132	130	183-187	185	183	
	O	9O	--	--	--	190-194	194	190-194	279	279	164-168	164-166	164	174-176	174-176	174-176	130-132	130-132	130	181	183	183	
	P	9P	--	--	--	194-200	190-200	190-200	279	279	166	158	166	192-194	192-194	192-194	130-132	130-132	130	183	187	187	
	Q	9Q	--	160	--	192-200	190-200	190-200	279	279	164	158-164	164	192-200	194-200	176-184	184-192	130-132	130	179-187	179-187	179-187	
	R	9R	--	160	--	200-202	194-200	194-200	260-279	260-279	164	164-166	164	164-166	176-192	176-192	176-192	130-132	130	185	183	183	
10	S	10S	152	200	--	200-208	200-208	208	279	279	162-164	164-168	162	162	192-196	174-184	174-192	130-132	130-132	130	181	185	
			152	200	--	200-208	200-208	208	279	279	162-164	164-168	162	162	192-196	174-184	174-192	130-132	130-132	130	181	185	

from male DNA, the Y-specificity of a microsatellite should be accepted. Here, we report that both markers do not meet with this 'rule of thumb' and phylogenetic studies in cattle should therefore be questioned.

This is the first time in which amplification of the microsatellite *INRA124* on female DNA is reported (Fig. S1). Moreover, even though this marker has previously only been found to be polymorphic between *B. taurus* and *B. indicus* (Hanotte *et al.* 1997, 2000; Edwards *et al.* 2007; Li *et al.* 2007), we have also found polymorphism within both *B. taurus* and *B. indicus*. This was checked by genotyping 200 taurine male samples belonging to 37 European cattle breeds, four Brahman, two Brangus and four Indian Zebu sire samples (Table S1). Three different alleles (130, 132 and 134 bp) were found in the taurine samples and other three (126, 128 and 132) on the zebu males.

*INRA124* has been frequently used in studies assessing admixture between *B. taurus* and *B. indicus* (Hanotte *et al.* 2000; Edwards *et al.* 2007; Li *et al.* 2007). As this marker undoubtedly amplifies the same PCR product on female taurine and zebu DNA, the occurrence of recombination events cannot be rejected. Therefore, the studies assessing male-mediated introgression events carried out on the basis of this marker alone should be interpreted with caution.

Some doubts on the Y-specificity of the marker *INRA126* already existed. PCR products of the same size from both sexes have been reported on cattle (Liu *et al.* 2003), yaks (Edwards *et al.* 2000) and saola (Nguyen *et al.* 2005). This could indicate that, across bovines, the X-chromosome has retained a homologous sequence to the Y-chromosomal segment containing the *INRA126* marker (Edwards *et al.* 2000) or that the location of the locus *INRA126* is on the PAR-BTAY (Li *et al.* 2007). Here, we show that *INRA126* amplifies from female DNA but also that this locus has a large variability within family. The large number of paternal incompatibilities that have accumulated on locus *INRA126* (see Table 2) cannot be easily explained, even though unexpectedly high mutation rates together with chromosomal recombination could be the basis of the large variability assessed on this locus (Li *et al.* 2007). Whatever the explanation might be, the use of the locus *INRA126* to characterize male evolutionary histories is questionable and not recommended.

Recently, Anderung *et al.* (2007) indicated the occurrence of hybridization between indicine and taurine cattle in Africa, and the presence of indicine X chromosomal recombinant forms. From our results, recombination events involving loci *INRA124* and *INRA126* cannot be rejected, and therefore use of these markers may distort actual indicine male-mediated ancestries.

Based on our findings, it is recommended that the *INRA124* and *INRA126* markers should not be used in studies to characterize male-mediated genetic events. Assessment of paternal genetic legacies would require thorough examination of BTAY-specific marker amplifica-

tion patterns and segregation on female DNA. Previous cattle population studies using these markers alone should be interpreted with caution. New analyses using Y-specific markers are necessary to gain a better understanding of the different levels of admixture in breeds from different geographic areas and the ages of the admixtures (Anderung *et al.* 2007).

## Acknowledgements

This study was partially funded by MEC-FEDER: CGL2005-03761/BOS and USDA-Hatch project MIN-16-019. L. Pérez-Pardal is funded by BES-2006-13545. We acknowledge Ligia Mercedes Jimenez Robayo, Maria Raquel Carvalho and Catarina Ginja for their help in sampling.

## References

- Anderung C., Hellborg L., Seddon J., Hanotte O. & Götherström A. (2007) Investigation of X- and Y-specific single nucleotide polymorphisms in taurine (*Bos taurus*) and indicine (*Bos indicus*) cattle. *Animal Genetics* 38, 595–600.
- Edwards C.J., Gaillard C., Bradley D.G. & MacHugh D.E. (2000) Y-specific microsatellite polymorphism in a range of bovid species. *Animal Genetics* 31, 127–30.
- Edwards C.J., Baird J.F. & MacHugh D.E. (2007) Taurine and zebu admixture in Near Eastern cattle: a comparison of mitochondrial, autosomal and Y-chromosomal data. *Animal Genetics* 38, 520–4.
- Giovambattista G., Ripoli M.V., De Luca J.C., Mirol P.M., Lirón J.P. & Dulout F.N. (2000) Male-mediated introgression of *Bos indicus* genes into Argentine and Bolivian Creole cattle breeds. *Animal Genetics* 31, 302–5.
- Hanotte O., Okomo M., Verjee Y., Rege E. & Teale A. (1997) A polymorphic Y chromosomal microsatellite locus in cattle. *Animal Genetics* 28, 308–22.
- Hanotte O., Tawah C.L., Bradley D.G., Okomo M., Verjee Y., Ocheng J. & Rege J.E. (2000) Geographic distribution and frequency of a taurine *Bos taurus* and an indicine *Bos indicus* Y specific allele amongst sub-Saharan African cattle breeds. *Molecular Ecology* 9, 387–96.
- Kappes S.M., Keele J.W., Stone R.T., McGraw R.A., Sonstegard T.S., Smith T.P., Lopez-Corrales N.L. & Beattie C.W. (1997) A second-generation linkage map of the bovine genome. *Genome Research* 7, 235–49.
- Li M.H., Zerabruk M., Vangen O., Olsaker I. & Kantanen J. (2007) Reduced genetic structure of north Ethiopian cattle revealed by Y-chromosome analysis. *Heredity* 98, 214–8.
- Liu W.-S. & Ponce de León F.A. (2007) Mapping of the bovine Y chromosome. *Electronic Journal of Biology* 3, 5–12.
- Liu W.-S., Beattie C.W. & Ponce de León F.A. (2003) Bovine Y chromosome microsatellite polymorphism. *Cytogenetic and Genome Research* 10, 53–8.
- MacHugh D.E. & Bradley D.G. (2001) Livestock genetic origins: goats buck the trend. *Proceedings of the National Academy of Sciences of the United States of America* 98, 5382–4.
- Nguyen T.T., Ménétrey F., Genini S., Nguyen V.L., Vögeli P., Nguyen B.X. & Stranzinger G. (2005) Application of bovine

564 Pérez-Pardal *et al.*

microsatellite markers on Saola (*Pseudoryx nghetinhensis*). *Journal of Animal Breeding and Genetics* **122**, 195–8.

Vaiman D., Imam-Ghali M., Moazami-Goudarzi K., Guerin G., Grohs C., Leveziel H. & Saidi-Mehdar N. (1994) Conservation of a syntenic group of microsatellite loci between cattle and sheep. *Mammalian Genome* **5**, 310–4.

### Supporting information

Additional supporting information may be found in the online version of this article.

**Figure S1** Genmapper image of *INRA124* amplification (green lines due to VIC labelling of PCR products). Bands of different size form the GenScan 500-LIZ standard (orange).

**Table S1** Number of genotyped individuals (N) and allele frequency found on locus *INRA124* in 37 bovine cattle breeds and some Zebu cattle samples.

**Appendix S1** PCR conditions and DNA sequences of the markers used.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting information supplied by the authors.







## Y-specific microsatellites reveal an African subfamily in taurine (*Bos taurus*) cattle

L. Pérez-Pardal<sup>\*,†</sup>, L. J. Royo<sup>\*</sup>, A. Beja-Pereira<sup>†</sup>, I. Curik<sup>‡</sup>, A. Traoré<sup>§</sup>, I. Fernández<sup>\*</sup>, J. Sölkner<sup>||</sup>, J. Alonso<sup>\*\*,†</sup>, I. Álvarez<sup>\*</sup>, R. Bozzi<sup>††</sup>, S. Chen<sup>†</sup>, F. A. Ponce de León<sup>††</sup> and F. Goyache<sup>\*</sup>

<sup>\*</sup>Área de Genética y Reproducción Animal, SERIDA, Camino de los Claveles 604 33203-Gijón, Spain. <sup>†</sup>CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Universidade do Porto, Campus Agrário de Vairão, 4485-661 Vairão, Portugal. <sup>‡</sup>Animal Science Department, University of Zagreb, Svetosimunska, Zagreb, Croatia. <sup>§</sup>INERA, 04 BP 8645 Ouagadougou 04, Burkina Faso. <sup>||</sup>BOKU – University of Natural Resources and Applied Life Sciences Vienna, Vienna, Austria. <sup>\*\*</sup>Centro de Inteligencia Artificial, Universidad de Oviedo, Campus de Viseques, 33271-Gijón, Spain. <sup>††</sup>Dipartimento di Scienze Zootecniche, Università degli Studi di Firenze, Via delle Cascine 5, 50144 Firenze, Italy. <sup>†††</sup>Department of Animal Science, University of Minnesota, St. Paul, MN, 55108, USA

### Summary

Five cattle Y-specific microsatellites, totalling six loci, were selected from a set of 44 markers and genotyped on 608 *Bos taurus* males belonging to 45 cattle populations from Europe and Africa. A total of 38 haplotypes were identified. Haplogroups (Y1 and Y2) previously defined using single nucleotide polymorphisms did not share haplotypes. Nine of the 27 Y2-haplotypes were only present in African cattle. Network and correspondence analyses showed that this African-specific subfamily clustered separately from the main Y2-subfamily and the Y1 haplotypes. Within-breed genetic variability was generally low, with most breeds (78%) showing haplotypes belonging to a single haplogroup. AMOVA analysis showed that partitioning of genetic variation among breeds can be mainly explained by their geographical and haplogroup assignment. Between-breed genetic variability summarized via Principal Component Analysis allowed the identification of three principal components explaining 94.2% of the available information. Projection of principal components on geographical maps illustrated that cattle populations located in mainland Europe, the three European Peninsulas and Mediterranean Africa presented similar genetic variation, whereas those breeds from Atlantic Europe and British Islands (mainly carrying Y1 haplotypes) and those from Sub-Saharan Africa (belonging to Y2-haplogroup) showed genetic variation of a different origin. Our study confirmed the existence of two large Y-chromosome lineages (Y1 and Y2) in taurine cattle. However, Y-specific microsatellites increased analytical resolution and allowed at least two different Y2-haplotypic subfamilies to be distinguished, one of them restricted to the African continent.

**Keywords** cattle, domestication, phylogeny, Y-chromosome, Y-specific markers.

### Introduction

Putative microsatellites mapping to the male-specific Y-chromosome (MSY) region of cattle are available (Bishop *et al.* 1994; Vaiman *et al.* 1994; Mezzelani *et al.* 1995; Kappes *et al.* 1997; Liu *et al.* 2003). Y-specific microsatellites have been used to trace cattle domestication and differentiation among bovid species (Edwards *et al.* 2000,

2007a). Some of these markers were also used to distinguish between taurine (*Bos Taurus*) and zebu (*B. indicus*) patriline and to infer introgression events from zebu into taurine populations (Bradley *et al.* 1994; Edwards *et al.* 2000, 2007a; Giovambattista *et al.* 2000; Hanotte *et al.* 2000; Li *et al.* 2007). Götherström *et al.* (2005) identified two Y-specific single nucleotide polymorphisms (SNPs) allowing two major clusters to be distinguished in taurine (*B. taurus*) cattle (Y1 and Y2). The frequencies of the identified SNPs in modern samples revealed that haplogroup Y1 can primarily be found in extant northern Europe cattle breeds and haplogroup Y2 is mainly distributed in the Central and Southern Europe cattle breeds. This strong geographical signal could be explained by a significant

Address for correspondence

Goyache, Área de Genética y Reproducción Animal, SERIDA, Camino de los Claveles 604 33203-Gijón, Spain.  
E-mail: goyache@serida.org

Accepted for publication 10 September 2009

influence of male aurochs on domesticated populations coming from the Near East during the formation of the present European cattle stock. This would challenge the common view on early domestication and Neolithic stock-rearing. However, Bollongino *et al.* (2008) and Svensson & Götherström (2008), using ancient and mediaeval samples, did not support the hypothesis that these markers distinguish European aurochs from domesticated cattle.

Y-specific microsatellites and SNPs have been jointly used to ascertain genetic history of human populations (Zerjal *et al.* 2002). SNPs provide an opportunity to understand paternal population origins, relationships and dispersals with greater phylogenetic and geographical resolution and less terminological ambiguity than was hitherto possible. The inclusion of microsatellite data allows the assessment of patterns of migration, admixture and ancestry, as well as to identify additional micro-evolutionary processes associated with population structure *sensu lato* (Zegura *et al.* 2004). Recently, extended Y-chromosome haplotypes that combine information from microsatellites and SNPs have been used to study genetic relationships among Portuguese (Ginja *et al.* 2009), Ethiopian (Li *et al.* 2007) and Eurasian (Kantanev *et al.* 2009) cattle populations. However, samples used are limited to these geographical areas.

The aim of this work was to investigate the haplotype structure and diversity in taurine cattle breeds from Europe and Africa using a combination of already described SNPs and microsatellites tested to be specific to the MSY region. The genetic relationships between breeds will also be analysed.

## Materials and methods

### Samples

DNA samples from 608 *B. taurus* males belonging to 45 cattle populations from Europe and Africa were analysed (Table 1). European sampling includes major breeds located in: (i) Atlantic Europe (12 breeds and 198 samples) including British, English Channel (including Normande) and European Friesian breeds; (ii) Continental Europe, including French (except Normande), German (except for European Friesian samples) and Austrian cattle breeds (13 breeds and 181 samples); (iii) Iberian Peninsula (16 breeds and 266 samples); (iv) Italian Peninsula (three breeds and 45 samples); and Balkan Peninsula (two breeds and 17 samples). African samples included two breeds from Mediterranean Africa (obtained in Morocco) totalling 10 samples and two Sub-Saharan Africa breeds (Lobi from Burkina Faso and N'Dama originated in Guinea) totalling 17 samples. Further description of sampling is given in Table 1.

### SNP analyses

Three polymorphisms previously described by Götherström *et al.* (2005) were analysed as follows: (i) polymorphism

DDX3Y1 (DBY), which differentiates zebu and taurine Y-specific haplogroups, was analysed on the African samples to ensure that further determinations were carried out on *B. taurus* Y-chromosomes; (ii) polymorphism *zfy10* was analysed on all samples to distinguish between haplogroups Y1 and Y2; and (iii) polymorphism *UTY19* was analysed on the African Y2 samples to confirm the haplogroup assignment. For SNP *UTY19* (C/A), PCR conditions were as follows: the final reaction volume was 10 µl; the reaction mixture consisted on 50–100 ng of genomic DNA, 0.2 U of Taq polymerase (Biotools), 0.25 µM of each primer (Sigma-Aldrich), 200 µM of each dNTP (Biotools) and 2 mM MgCl<sub>2</sub>. The PCR protocol included an initial step of 95 °C (3 min), followed by 40 cycles of 30 s at 95 °C for DNA denaturation, 30 s for primer annealing at 54 °C and 30 s at 72 °C for primer extension. PCR product was purified with the exoSAP-IT<sup>®</sup> protocol (USB) and sequenced in both strands by using the BigDye<sup>®</sup> Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) following the manufacturer's recommendations, on a ABI3130 sequence analyser (Applied Biosystems). A DX3Y1 C/T polymorphism (AY928811:g.437) differentiates taurine and zebu Y-chromosome, and an UTY G/T polymorphism (AY936543:g.423) differentiates Y1 and Y2 haplogroup. For the *ZFY10* indel, the PCR protocol was as follows: the final reaction volume was 10 µl; the reaction mixture consisted of 50–100 ng of genomic DNA, 0.2 units of Taq polymerase (Biotools), 0.25 µM of each primer (Sigma-Aldrich), 200 µM of each dNTP (Biotools) and 2 mM MgCl<sub>2</sub>. The PCR protocol included an initial step of 95 °C (3 min), followed by 30 cycles of 30 s at 95 °C for DNA denaturation, 30 s for primer annealing at 55 °C and 30 s at 72 °C for primer extension. An additional final step of 72 °C for 90 min was included to avoid n-1 band presence when using capillary electrophoresis (Table 2). Genotyping was carried out by direct electrophoreses on an ABI3130 sequencer analyser (Applied Biosystems) based on its difference in size: 285–287 nt for Y1 and Y2 respectively.

### Microsatellite genotyping

A total of 43 microsatellites previously reported to be located on the bovine MSY region by Matthews *et al.* (1991), Bishop *et al.* (1994), Vaiman *et al.* (1994), Mezzelani *et al.* (1995), Kappes *et al.* (1997), Liu *et al.* (2003), and microsatellite *SRM18* described on the ovine Y-chromosome by Meadows *et al.* (2006), were tested for male specificity, paternal compatibility and repeatability scoring following the protocol described in Pérez-Pardal *et al.* (2009; see Table S1). After this selection process only five microsatellites were found to be suitable for our study: *INRA189*, *UMN0103*, *UMN0307*, *BM861* and *BYM1*. Microsatellite *UMN0103* showed a bimodal allele-frequency distribution that could be scored as two different loci

**Table 1** Distribution of the 608 male individuals analysed per breed and country. Each breed is assigned to a given geographical area: (i) Atlantic Europe; (ii) Continental Europe; (iii) Balkan Peninsula; (iv) Iberian Peninsula; (v) Italian Peninsula; (vi) Mediterranean Africa; and (vii) Sub-Saharan Africa. Assignment of samples to each identified haplogroup is also given (number of Y2 samples belonging to the African Y2 haplotypes are in brackets). Expected heterozygosity (He), haplotype diversity corrected for sample size (h) and standard errors (SE) are also given.

Breed	Country	Sample size	Geographical areas	Y1	Y2	He	n	h	SE(h)
Alistana	Spain	9	Iberian Peninsula		9	0	1	0.125	0.022
Angus	UK	13	Atlantic Europe	13		0.083	2	0.167	0.030
Asturiana de la Montaña	Spain	19	Iberian Peninsula	18	1	0	1	0.056	0.002
Asturiana de los Valles	Spain	38	Iberian Peninsula	32	6	0.321	3	0.081	0.003
Blanc-Bleu Belge	Belgium	15	Atlantic Europe	15		0.116	2	0.143	0.021
Berrenda	Spain	4	Iberian Peninsula		5	0.213	3	0.750	0.344
Betizu	Spain	10	Iberian Peninsula		10	0	1	0.111	0.016
Blonde d'Aquitaine	France	5	Continental Europe		5	0.133	3	0.750	0.344
Brown Swiss	Switzerland	14	Continental Europe		14	0	1	0.077	0.004
Bruna des Pirineus	Spain	11	Iberian Peninsula		11	0.066	2	0.200	0.043
Bovin d'Oulmes	Morocco	5	Mediterranean Africa		5	0.053	2	0.500	0.219
Bovin de Tidili	Morocco	5	Mediterranean Africa		5 (1)	0.347	2	0.500	0.219
Charolais	France	5	Continental Europe		5	0	1	0.250	0.094
Chianina	Italy	20	Italian Peninsula		20	0	1	0.053	0.002
Croatian Simmental	Croatia	13	Balkan Peninsula	3	10	0.237	2	0.167	0.030
Danish Red	Denmark	10	Atlantic Europe	9	1	0.120	2	0.222	0.053
Dexter	Ireland	4	Atlantic Europe	4		0	1	0.333	0.156
Fleckvieh	Switzerland	4	Continental Europe		4	0	1	0.333	0.156
European Friesian	Germany <sup>1</sup>	8	Atlantic Europe		8	0.036	2	0.286	0.084
Galloway	UK	2	Atlantic Europe		2	0.083	2	2.000	1.937
Hereford	UK	8	Atlantic Europe		7	0.146	2	0.286	0.084
Holstein	The Netherlands	37	Atlantic Europe	37		0	1	0.028	0.006
Istrian	Croatia	4	Balkan Peninsula		4	0	1	0.333	0.156
Jersey	The Netherlands	5	Atlantic Europe		5	0	1	0.250	0.094
Toro de Lidia	Spain	65	Iberian Peninsula		65	0.084	3	0.047	0.002
Limousin	France	4	Continental Europe		4	0.063	2	0.667	0.349
Lobi	Burkina Faso	8	Sub-Saharan Africa		8 (8)	0.307	1	0.143	0.030
Maremmna	Italy	19	Italian Peninsula		19	0	1	0.056	0.002
Morucha	Spain	17	Iberian Peninsula		17	0.219	3	0.188	0.031
N'Dama	Guinea <sup>2</sup>	9	Sub-Saharan Africa	2	7 (7)	0.300	3	0.375	0.110
Normande	France	39	Atlantic Europe	39		0	1	0.026	0.006
Pajuna	Spain	4	Iberian Peninsula	1	3	0.354	3	1.000	0.541
Parda de Montaña	Spain	26	Iberian Peninsula	1	25	0.049	2	0.080	0.003
Parthenais	France	15	Continental Europe	4	11	0.261	2	0.143	0.021
Piemontese	Italy	6	Italian Peninsula		6	0.185	3	0.600	0.240
Pinzgauer	Austria	9	Continental Europe		9	0	1	0.125	0.022
Pirenaica	Spain	10	Iberian Peninsula		10	0.053	2	0.222	0.053
Retinta	Spain	6	Iberian Peninsula		6	0.148	2	0.400	0.151
Rubia Gallega	Spain	31	Iberian Peninsula		31	0	1	0.033	0.005
Salers	France	15	Continental Europe	1	14	0.083	2	0.143	0.021
Sayaguesa	Spain	4	Iberian Peninsula		4	0.125	3	1.000	0.541
Serrana Negra	Spain	3	Iberian Peninsula		3	0.148	2	1.000	0.663
Simmenthal	Germany	4	Continental Europe		4	0	1	0.333	0.156
Tarentaise	France	18	Continental Europe		18	0	1	0.059	0.001
Tudanca	Spain	8	Iberian Peninsula		8	0	1	0.143	0.030
Tyrolean Grey	Austria	19	Continental Europe		19	0.031	2	0.111	0.011
Totals				202	406 (16)		38	0.063	0.000

<sup>1</sup>These samples were mainly from German sires. However, sires from Sweden and The Netherlands are included in this group.

<sup>2</sup>N'Dama samples, obtained in Congo from herds managed by the Compagnie JVL, have been assigned to Guinea because of the historical origin of the animals forming the sampled herds.

**Table 2** Description of the Y-specific microsatellites and SNPs typed in this work. The microsatellites *INRA189*, *UMN0103*, *UMN0307*, *BM861* and *BYM1* were analysed in all samples. Two multiplex PCRs reactions (PCR 1: *INRA189* and *UMN0103* and PCR 2: *UMN0307*, *BM861* and *BYM1*) were carried out. Polymorphism *ZFY10* was analysed on all samples to distinguish between haplogroups Y1 and Y2. SNP *UTY19*, which differentiates zebu and taurine Y-specific haplogroups, was analysed on the African samples to ensure that further determinations are carried out on *Bos taurus* Y-chromosomes.

Microsatellite	Repeat pattern	Primer sequence	Primer concentration ( $\mu$ M)	Temperature ( $^{\circ}$ C)	Labelling	No. of alleles
INRA189	(TC) <sub>n</sub>	TTTGTTCCTGCTGAG-F	0.25	58	VIC	8
		GAACCTCGTCTCTGTAGCC-R	0.25		No	
UMN0103	(CA) <sub>n</sub>	ACACAGAGTATTCACCTGAG-F	0.25	58	FAM	9
		ATTTACCTCGGTCAAAGCAC-R	0.25		No	
UMN0307	(CA) <sub>n</sub>	GATACAGCTGAGTACTAAC-F	0.18	58	PET	5
		GTGCACATCTGAGCTGTG-R	0.18		No	
BYM-1	(AC) <sub>n</sub>	CCTTGTTGACCTTGACCACT-F	0.18	58	FAM	3
		TTCCAGGCACAGAAACGGA-R	0.18		No	
BM861	(GT) <sub>n</sub> (CT) <sub>n</sub>	TTGAGCCACTCGAAAGC-F	0.18	58	FAM	2
		CAAGCGTGTGTTGATG-R	0.18		No	
DDX3Y1	-	CGAAACACAGAACTTAATCT-F	0.25	54	No	2
		AGTATGTCATTAGCCATATC-R	0.25		No	
UTY 19	-	CGTAGCAACACCTCTTTCTG-F	0.25	54	No	2
		CAAGCTTCAAGTTGTTAC-R	0.25		No	
ZFY10	-	CCAAATCGTTGAGCTTATGA-F	0.25	55	FAM	2
		CGAGCATAAGTGATCCAATGAA-R	0.25		No	

(Balaesque *et al.* 2006). Therefore the five selected microsatellites included six different loci.

The five selected microsatellites were analysed on all samples in two multiplex PCRs (PCR 1: *INRA189* and *UMN0103* and PCR 2: *UMN0307*, *BM861* and *BYM1*) performed in a GenAmp PCR system 9700 (Applied Biosystems). The final reaction volume was 10  $\mu$ l. The reaction mixture consisted of 50–100 ng of genomic DNA, 0.2 units of Taq polymerase (Biotools), 200  $\mu$ M of each dNTP (Biotools) and 2 mM  $MgCl_2$ . Concentration primers (Sigma-Aldrich) and fluorescent dye labels are given in Table 2. The PCR protocol included an initial step of 95  $^{\circ}$ C (3 min), followed by 30 cycles of 30 s at 95  $^{\circ}$ C for DNA denaturation, 30 s for primer annealing at 58  $^{\circ}$ C and 30 s at 72  $^{\circ}$ C for primer extension. An additional final step of 72  $^{\circ}$ C for 90 min was included to avoid n-1 band presence when using capillary electrophoresis. PCR fragments were separated by capillary electrophoresis on an ABI 310 instrument (Applied Biosystems) according to the manufacturer's recommendations. Allele sizes were determined with Genemapper (Applied Biosystems).

### Terminology

The terms 'haplogroup' and 'haplotype' have various, overlapping definitions in the literature. Throughout the article, we will use the terminology proposed by de Knijff (2000) and recommended by the Y-Chromosome Consortium (2002) in which 'haplogroup' refers to MSY lineages defined by binary polymorphisms and the term 'haplotype' is

reserved for all subfamilies of haplogroups that are defined by variation at microsatellites on the MSY region.

The two different loci shown by microsatellite *UMN0103* will be analysed using the Mathias's prior (Balaesque *et al.* 2006). Mathias *et al.* (1994) assigned the co-amplified alleles to each locus according to their sizes, the shortest allele being assigned to locus 1, the longest to locus 2.

### Statistical analyses

Microsatellite alleles were combined into haplotypes. Observed haplotypes were analysed as follows: (i) analysis of correspondence was performed using the Proc CORRESP of the SAS/STAT package (SAS Institute Inc.); and (ii) a median-joining network connecting different haplotypes was constructed using the program NETWORK 4.5.2 [available at <http://www.fluxus-engineering.com/> (Bandelt *et al.* 1999)]. To avoid reticulation, a reduced median algorithm (Bandelt *et al.* 1995) was used to generate a .rmf file and the median joining network method (Bandelt *et al.* 1999) was applied to this file.

An unbiased estimate of haplotype diversity,  $h$ , and its variance,  $V(h)$ , were calculated according to the method of Nei (1987, formulas 8.5 and 8.13 therein). The standard error of  $h$ ,  $SE(h)$ , was calculated by taking the square root of  $V(h)$ . Single-locus gene-diversity values were calculated in the same way. Within-breed expected heterozygosity adjusted for sampling size was also computed according to Nei (1987). The between-breeds genetic identity matrix was computed, according to Nei (1987), as the average across loci of the term to

$$\sum_i x_{ij} y_{ij}$$

where  $x_{ij}$  and  $y_{ij}$  are the frequencies of the  $i$ th allele at the  $j$ th locus within the populations  $x$  and  $y$  using the program MOLKIN (Gutiérrez *et al.* 2005). Information provided by this matrix was summarized by conducting principal component analyses (PCA) on the complement of this matrix. The principal component scores were used to construct interpolation maps drawn using the SPATIAL ANALYST EXTENSION OF ARCVIEW, available at: <http://www.esri.com/software/arcview/>. The Inverse Distance Weighted (IDW) option with a power of two was selected for the interpolation of the surface. IDW assumes that each input point has a local influence that diminishes with distance. The area of sampling of each breed was used as geographical coordinates, and the six nearest neighbours were used for the calculation. Interpolation surfaces were divided into eight equal classes.

AMOVA analyses were computed using the program ARLEQUIN 3.11 (Excoffier *et al.* 2005). Analyses were carried out fitting one single (haplogroup, haplofamily, breed, continent or geographical area) or two hierarchical levels (breed within continent or breed within geographical area).

## Results

The number of alleles per polymorphic microsatellite varied from 2 (BM861) to 8 (INRA189), revealing a total of 32 alleles across the six loci used. Overall gene diversity was 0.429. Altogether, the identified alleles defined 38 haplotypes (overall haplotype diversity of  $0.063 \pm 0.0002$ ). Haplotype descriptions and their distribution across breeds are given in Table S2. Samples were assigned to the Y-chromosome haplogroups defined by Götherström *et al.* (2005) according to the SNPs typed. UTY19 polymorphism confirmed the Y2-haplogroup assignment of the African samples. A total of 11 haplotypes belonged to haplogroup Y1 (named from Y1\_1 to Y1\_11) whereas 27 haplotypes (from Y2\_1 to Y2\_27) belonged to haplogroup Y2.

Haplogroups did not share haplotypes. The most common Y1-haplotype (Y1\_6) was identified in 106 samples belonging to seven different breeds. The most common Y2-haplotype (Y2\_19) was identified in 269 samples across 29 breeds. Three Y1-haplotypes and eleven Y2-haplotypes were unique. Haplotype Y1\_6 is basically fixed in the Normande, Holstein and European Friesian breeds. Haplotype Y2\_19 was widely distributed across breeds, being dominant in Central and Southern European and Moroccan samples. A group of nine Y2-haplotypes (Y2\_1, Y2\_3, Y2\_4, Y2\_5, Y2\_7, Y2\_9, Y2\_10, Y2\_11 and Y2\_27) were only present in Sub-Saharan cattle but also in one Moroccan sample (haplotype Y2\_1).

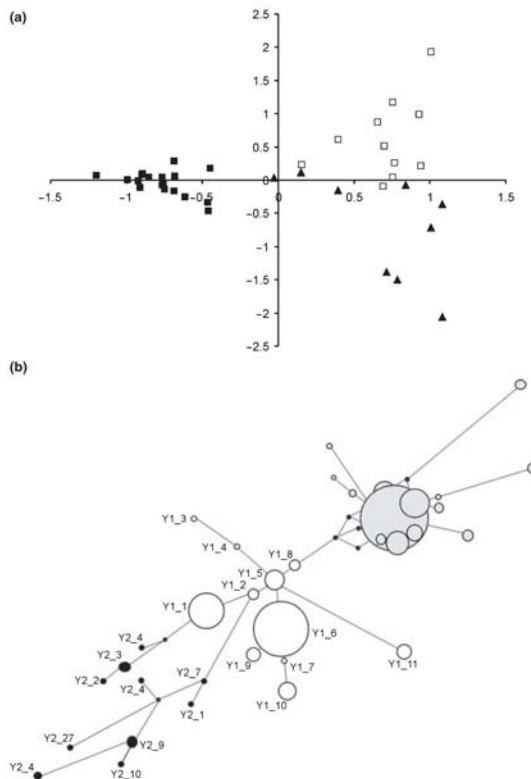
Figure 1 summarizes the genetic variability contained by the 38 haplotypes identified. Correspondence analysis separated the main group of Y2-haplotypes from the Y1-haplotypes on Dimension 1 (X-axis). Interestingly

enough, the subfamily formed by the nine African specific Y2-haplotypes were also separated from the main Y2-subfamily on the X-axis; differentiation between this second Y2-subfamily and the Y1 haplotypes was assessed on Dimension 2 (Y-axis). The existence of 2 subfamilies within the Y2-haplogroup was confirmed via NETWORK analysis. The Y2-subfamily identified on African sire Y-chromosomes clustered with the Y1-haplotypes.

Within-breed genetic variability was generally low (Table 1). Most breeds (78%) showed haplotypes belonging to a single haplogroup. Seventeen (37%) of the analysed breeds showed a single haplotype. Within-breed expected heterozygosity varied from 0 to 0.354 (Pajuna breed). Within-breed haplotype diversity showed more variation, ranging from  $0.026 \pm 0.006$  in the Normande breed to 1.0 in some endangered local breeds such as Sayaguesa, Serrana Negra and Pajuna breeds. In general, breeds that are considered closely related, such as those belonging to Brown Swiss type (Brown Swiss, Tyrolean Grey, Parda de Montaña and Bruna dels Pirineus) or the classical European Friesian and the modern Holstein cattle, have consistently the same haplotype composition, the most frequent being Y2\_19 and Y1\_6 respectively. Between-breed genetic relationships were summarized via PCA performed on the complement of the between-breeds genetic identity matrix (Table S3). Three principal components (PC) with eigenvalue  $>1$ , explaining 81.8%, 10.2% and 2.2% of the total variability, were retained (Table S2). PC1 included mainly the Central European, Southern European and Moroccan breeds; PC2 mainly included Atlantic European breeds; PC3 was almost exclusively formed by Sub-Saharan Lobi and N'Dama breeds. This pattern was consistent with the variability previously shown by the identified haplotypes. Projection of PCs on a geographical map is shown in Fig. 2. The three PCs gave very similar images: cattle populations located in mainland Europe, the three European Peninsulas and Mediterranean Africa presented similar genetic variation whereas those breeds from Atlantic Europe and British Islands (mainly carrying Y1 haplotypes) and those from Sub-Saharan Africa (belonging to Y2-haplogroup) showed genetic variation of different origin.

Genetic identity between Sub-Saharan Africa populations and most other populations (regardless they are formed by 'European' Y2 samples) was below 0.30. This is only comparable with the genetic identity assessed between the Atlantic European sample and the other geographical cattle populations, ranging from 0.322 to 0.426 (Table 3). Samples obtained from breeds of the mainland European continent, the three European Peninsulas and Mediterranean Africa have between-population genetic identities ranging from 0.664 to 0.873. Atlantic European and Sub-Saharan samples have poor genetic identity (0.315).

This general scenario was confirmed by an AMOVA analysis (Table 4). The 'Continent' only accounted for roughly 20% of the variation and 'Geographical Area' explained less than 50% of the variation. When 'Breed' was included in



**Figure 1** Graphical representations of genetic variability among Y1 and Y2 haplotypes. Plot a shows the dispersion of the observed 38 haplotypes calculated via correspondence analysis. Y1-haplotypes are in black triangles; Y2-haplotypes are in black squares except for those specific haplotypes found in African cattle, represented by open squares. Plot b shows a network tree constructed using the program NETWORK 4.5.2. The plot shows Y1-haplotypes (open circles), Y2-haplotypes (grey circles) except for those specific haplotypes for African cattle (black circles). Circles are proportional to haplotype frequencies. Regardless of the analysis carried out, the African-specific Y2-haplotypes tend to cluster nearer the Y1-haplotypes.

the model as the only hierarchical factor to be tested, it explained roughly 80% of the genetic variation. This variation was similar to the proportion of variation explained by 'Haplogroup' (71.09%) and 'Haplofamily' (74.51%). When 'Breed' was analysed within Continent or Geographical Area, it still accounted for more genetic variation than these latter two factors. The variation explained by the haplofamily was slightly (but significantly) higher than that explained by the haplogroup.

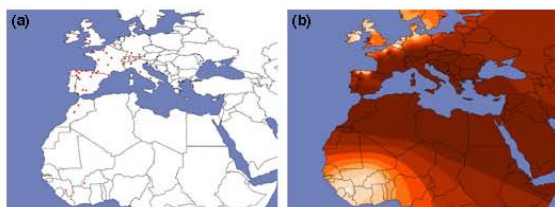
## Discussion

Here we use a set of five cattle Y-specific microsatellites, including six loci, to assess diversity on the male path in taurine cattle. The microsatellite set used was fitted after

careful selection of markers (see Table S1). Our microsatellites basically coincide with those previously used by Ginja *et al.* (2009) in Portuguese cattle, and three of them (INRA189, BM861 and BYM-1) were also used by Li *et al.* (2007) and Kantanen *et al.* (2009) in Ethiopian and Eurasian cattle respectively. The markers INRA124, INRA 126 and UMNO504 used in the aforementioned studies have been shown not to be male-specific here and in the previous analyses (Pérez-Pardal *et al.* 2009).

Haplotype diversity found in our sample is similar to that of 0.064 (26 haplotypes among 405 bulls tested) reported by Kantanen *et al.* (2009), higher than that of 0.042 (13/307) reported by Ginja *et al.* (2009) in Portuguese cattle, but lower than that of 0.099 (16/161) found by Li *et al.* (2007) in Ethiopian cattle (including Holstein-Friesian).





**Figure 2** Locations of cattle populations sampled (Plot a) and synthetic maps illustrating geographical variation of the first principal component (Plot b) identified using principal component analyses (PCA). The projection of the main component of variation represents the main Y2 haplotypic subfamily, predominant in the cattle breeds sampled in Central Europe, the three European Peninsulas and Mediterranean Africa. By contrast, the Y1 haplotypes are the most frequent in the Atlantic European coastline, and a second Y2 haplotypic subfamily is typical of the Sub-Saharan cattle.

The results reported by Li *et al.* (2007) may be biased upwards as a result of the presence of taurine and zebu haplotypes in the Ethiopian samples, but also by the use of the putative fast-evolving marker INRA126 that has been shown to have erratic behaviour with a high rate (55%) of parental incompatibility (Pérez-Pardal *et al.* 2009); it may be located on the pseudoautosomal region of the Y-chromosome (Li *et al.* 2007).

At the subspecies level, our study highlights the following: (i) haplogroups Y1 and Y2 previously identified by Götherström *et al.* (2005) have phylogenetic meaning; (ii) there is genetic variation within haplogroups that can be assessed using microsatellites; (iii) African cattle Y2-chromosomes can be differentiated from those of European cattle; (iv) there is clear geographical structuring in taurine cattle populations; and (v) this structuring is compatible with a poor overall genetic differentiation.

The fact that no haplotypes were shared between haplogroups clearly confirms the phylogenetic importance of the haplogroups identified by Götherström *et al.* (2005). They characterize two large cattle Y-chromosome lineages that are predominant in the European Atlantic coastline (Y1) and in the mainland European continent and Africa (Y2). However, in the case of cattle, biallelic markers, such as the Götherström *et al.* (2005) SNPs and indels, do not offer enough resolution to characterize Y-chromosome diversity. Y-specific microsatellites allowed at least two different Y2-lineages to be differentiated (Fig. 1). The African-specific

Y2-lineage clusters nearer to Y1 haplotypes than to European Y2-haplotypes. This suggests the existence of different wild sire populations with poor genetic differentiation. However, the set of markers used was not powerful enough to correctly discriminate among haplogroups. The identification of new Y-specific markers in cattle with higher resolution should confirm the genetic relationships between the different haplogroups.

Ginja *et al.* (2009) suggested a possible African introgression into Portuguese cattle based on one allele on marker INRA189 previously found in NDama cattle (Edwards *et al.* 2000). Differences between West African genetic stock and European Y2-sires are more marked than previously shown. Here we demonstrate that these breeds have their own haplotypes that are not shared with any other Y2-cattle breeds. Haplotype diversity, adjusted for sample size, is quite similar for Y1 and Y2 haplogroups ( $0.055 \pm 0.0006$  and  $0.067 \pm 0.0002$ , respectively). In any case, the higher haplotype diversity observed in Y2 cattle is mainly resulting from the variation observed in African Y2-haplotypes. The identified West African Y2-subfamily has an extremely high haplotype diversity  $0.533 \pm 0.1214$ , whereas the other Y2 samples had haplotype diversity of  $0.049 \pm 0.0007$ . Although our analyses do not allow us to reject the influence of other factors such as differences in population sizes or introgression from other populations, this fact suggests that West African cattle gather genetic variation of local origin. Summarizing the

Geographical area	1	2	3	4	5	6	7
1. Iberian Peninsula	0.629						
2. Atlantic Europe	0.363	0.775					
3. Continental Europe	0.736	0.353	0.913				
4. Italian Peninsula	0.704	0.328	0.873	0.892			
5. Balkan Peninsula	0.694	0.426	0.849	0.808	0.806		
6. Mediterranean Africa	0.664	0.322	0.814	0.780	0.756	0.753	
7. Sub-Saharan Africa	0.284	0.315	0.249	0.230	0.246	0.253	0.574

**Table 3** Within-population (on diagonal) and between-population genetic identities (below diagonal) calculated by geographical area grouping of samples.

**Table 4** Geographical structuring assessed using Analysis of Molecular Variance (AMOVA). AMOVA subdivides the genetic diversity into hierarchical components and estimates the indices  $\Phi$ , which are molecular equivalents of Wright's  $F$  statistics and can be interpreted as follows:  $\Phi_{ST}$ : relative divergence between populations;  $\Phi_{SC}$ : relative divergence between populations of the same group;  $\Phi_{CT}$ : relative divergence between groups. The statistical significance of variance components and  $\Phi$  indices is evaluated by bootstrapping using 1000 replications. All estimates were statistically significant for  $P < 0.0001$ .

Type of grouping	Percentage of variance components			Total Variance	$\Phi$ Statistics		
	Among groups	Among populations within groups	Within populations		$\Phi_{ST}$	$\Phi_{SC}$	$\Phi_{CT}$
Continents (Europe and Africa)	19.95	—	80.05	1.569	0.199	—	—
Geographical area <sup>1</sup>	46.95	—	53.05	1.486	0.469	—	—
Breed	79.76	—	20.24	1.328	0.798	—	—
Haplogroup (Y1 and Y2)	71.09	—	28.91	2.128	0.711	—	—
Haplofamily <sup>2</sup>	74.51	—	25.49	2.106	0.745	—	—
Breeds within continents	11.55	70.34	18.11	1.484	0.819	0.795	0.115
Breeds within geographical areas <sup>2</sup>	39.18	42.61	18.21	1.476	0.818	0.701	0.392

<sup>1</sup>Geographical areas fitted were as follows (see Materials and methods and Table S1): Atlantic Europe (including the UK breeds, Dexter, Danish Red, Jersey, European and Holstein Friesian, Blanc-Bleu Belge and Normande cattle breeds); Iberian Peninsula (including Spanish breeds); Italian Peninsula (including Piemontese, Maremmana and Chianina breeds); Balkan Peninsula (including Croatian Simmental and Istrian cattle breeds); European mainland (including the rest of the European cattle breeds); Mediterranean Africa (including Bovine d'Oulmes and Bovine de Tidj samples); and Sub-Saharan Africa (including Lobi and NDama cattle samples).

<sup>2</sup>The haplofamilies fitted here are: (i) Y1-haplotypes; (ii) African-specific Y2 haplotypes; and (iii) the remaining Y2 haplotypes.

available genetic and archaeological evidence, Bradley & Magee (2006) suggested that African taurine cattle are not a simple subset of that domesticated in the Near East. This would be supported by the qualitatively different nature of their mtDNA composition, showing a large proportion of unique haplotypes (Troy *et al.* 2001), and that there is a significant component of microsatellite variation within African cattle that seems to be indigenous in origin (Hanotte *et al.* 2002). Though sample size from West African cattle is limited (only 17 samples), the evidence summarized by Bradley & Magee (2006) can also be applied to the Y-chromosome scenario identified here. In this respect, our findings would be consistent with a putative domestication event in Africa, including the use of local Y2-like wild sires. Moreover, the presence of specific African Y-chromosome haplotypes was found neither in Iberian cattle nor in the cattle of the other Southern European Peninsulas. Beja-Pereira *et al.* (2006) found African T1 mtDNA haplotypes in Southern and Northern Iberian cattle breeds. The lack of presence of African Y-specific haplotypes in Iberian cattle does not support recent introgression events. Genetic relationships between Iberian and African cattle populations would predate the traditional route associated with the Moorish invasions and occupation (Beja-Pereira *et al.* 2006).

As previously shown by Götherström *et al.* (2005), the Y1 and Y2 haplogroups have a clear geographical structure, with Y1 samples basically restricted to the Atlantic Europe and British Islands. The analysed breeds usually show Y-chromosomes belonging to a single haplogroup and are, in general, homogeneous at the Y-chromosome variation level. In most cases, the presence of Y1 and Y2 haplotypes

in a given breed can be explained by its recent history. Examples are the process of introgression of Fleckvieh, Brown Swiss and Friesian cattle into the Asturian cattle breeds during the first half of the 20th century (Goyache 1995) or the presence of Y2-haplotypes in Danish Red cattle, which is known to be introgressed by Brown Swiss cattle (Sørensen *et al.* 2005); the Red and White (Friesian) introgressed into Simmental cattle in Croatia; and the Y1 haplotypes found in NDama samples may have originated from Devon cattle influencing the sampled population during the 1930s.

The genetic heterogeneity of the wild sires used to create modern European cattle is a question of much debate. Götherström *et al.* (2005) suggested a large genetic influence of the primitive European aurochs in the formation of the present cattle breeds in Europe. However, Edwards *et al.* (2007b) and Bollongino *et al.* (2008), using ancient DNA, suggested that the model of a rapid introduction into Central Europe of cattle domesticated in the Near East without significant crossbreeding with local wild cattle remains unchallenged. Svensson & Götherström (2008) also suggested that the geographical consistency of the Y-specific haplogroups identified in European cattle could simply be the result of drift processes that occurred just before and during the formation of modern cattle breeds. The current analysis cannot support the hypothesis proposed by Götherström *et al.* (2005) suggesting that haplogroup Y2 would result from the domestication events occurring in the Near East during the Neolithic period, whereas Y1 would result from introgression of local male aurochs into domestic cattle populations all over the Atlantic European coastline. Both the haplotypic diversity and the



within-population expected heterozygosity (the complementary value of genetic identity) across different parts of Europe (Table S2) are low and substantially the same. This dataset suggests that both European Y1 and Y2 haplotypic families are associated with small initial numbers of sires. This would be consistent with a single domestication process. This contrasts with the high haplotypic variability and relatively high expected heterozygosity assessed in West African samples, in which there may have been the use of local sires (through genetically related Y2 sires domesticated in the Near East). The lack of samples from the Near East (the only definite centre of taurine cattle domestication) does not allow us to check this hypothesis.

Although there is room for improvement, this has been the first attempt, to our knowledge, to ascertain Y-specific genetic diversity and structure in African and European taurine cattle. Y-specific microsatellites have provided information about differentiation between Y-haplotypic lineages and relationships between cattle populations and geographical areas. Increasing the number of markers and the extent to which their properties are understood will influence the reliability of the inferences. Further analyses, including of samples from the Near East, as well as from different African regions, will help to clarify the role of Africa in the domestication of taurine cattle.

### Acknowledgements

This work was partially supported by the MICIN project no. CGL2005-03761/BOS and the FCT grant POCI/CVT/56758/2004. LP-P is supported by grant MICINN BES-2006-13545. ABP and SC are supported by FCT grants SFRH/BPD/38096/2007 and SFRH/BPD/26802/2006 respectively. APdL is supported by USDA-Hatch project MIN-16-019. Authors thank Eduardo Miura (Unión de Criadores de Toros de Lidia, Madrid), Pedro J. Azor (Universidad de Córdoba, Spain), Juan José Arranz (Universidad de León, Spain), Dimitri Pirotin (Université de Liège), Davide de Petris (Università degli Studi di Firenze, Italy), Jordi Jordana (Universitat Autònoma de Barcelona), Thomas Lecomte (JVL Company, Belgium-Congo) and Beate Scherf (FAO, Rome, Italy) for their support and help with sampling.

### References

Balaresque P., Sibert A., Heyer E. & Crouau-Roy B. (2006) Unbiased interpretation of haplotypes at duplicated microsatellites. *Annals of Human Genetics* **71**, 209–19.

Bandelt H.-J., Forster P., Sykes B.C. & Richards M.B. (1995) Mitochondrial portraits of human populations. *Genetics* **141**, 743–53.

Bandelt H.-J., Forster P. & Röhl A. (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* **16**, 37–48.

Beja-Pereira A., Caramelli D., Lalueza-Fox C. et al. (2006) The origin of European cattle: evidence from modern and ancient

DNA. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 8113–8.

Bishop M.D., Kappes S.M., Keele J.W. et al. (1994) A genetic linkage map for cattle. *Genetics* **136**, 619–39.

Bollongino R., Elsner J., Vigne J.-D. & Burger J. (2008). Y-SNPs do not indicate hybridisation between European Aurochs and domestic cattle. *PLoS One*, **3**, <http://dx.doi.org/10.1371/journal.pone.0003418>.

Bradley D.G. & Magee D.A. (2006) Genetics and origins of domestic cattle. In: *Documenting Domestication. New Genetic and Archaeological Paradigms* (Ed. by M.A. Zeder, D.G. Bradley, E. Emshwiller & B.D. Smith), pp. 317–28. University of California Press, Berkeley-Los Angeles-London.

Bradley D.G., MacHugh D.E., Loftus R.T., Sow R.S., Hoste C.H. & Cunningham E.P. (1994) Zebu-aurine variation in Y chromosome DNA: a sensitive assay for genetic introgression in West African trypanotolerant cattle populations. *Animal Genetics* **25**, 7–12.

Edwards C.J., Gaillard C., Bradley D.G. & MacHugh D.E. (2000) Y-specific microsatellite polymorphism in a range of bovid species. *Animal Genetics* **31**, 127–30.

Edwards C.J., Baird J.F. & MacHugh D.E. (2007a) Taurine and zebu admixture in Near Eastern cattle: a comparison of mitochondrial, autosomal and Y-chromosomal data. *Animal Genetics* **38**, 520–4.

Edwards C.J., Bollongino R., Scheu A. et al. (2007b) Mitochondrial DNA analysis shows a Near Eastern Neolithic origin for domestic cattle and no indication of domestication of European aurochs. *Proceedings Biological Sciences* **274**, 377–385.

Excoffier L., Laval G. & Schneider S. (2005) Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1**, 47–50.

Ginja C., Telo da Gama L. & Penedo M.C.T. (2009) Y Chromosome haplotype analysis in Portuguese cattle breeds using SNPs and STRs. *Journal of Heredity* **100**, 148–57.

Giovambattista G., Ripoli M.V., De Luca J.C., Mirol P.M., Lirón J.P. & Dulout F.N. (2000) Male-mediated introgression of *Bos indicus* genes into Argentine and Bolivian Creole cattle breeds. *Animal Genetics* **31**, 302–5.

Götherström A., Anderung C., Hellborg L., Elburg R., Smith C., Bradley D.G. & Ellegren H. (2005) Cattle domestication in the near East was followed by hybridization with aurochs bulls in Europe. *Proceedings of the Royal Society B* **272**, 2345–50.

Goyache F. (1995) *Influencia de efectos sistemáticos sobre los caracteres de importancia económica en la raza Asturiana de los Valles*. PhD Thesis. Universidad Complutense de Madrid, Madrid, Spain. ISBN: 84-8466-997-1 (2002).

Gutiérrez J.P., Royo L.J., Álvarez I. & Goyache F. (2005) MolKin v2.0: a computer program for genetic analysis of populations using molecular coancestry information. *Journal of Heredity* **96**, 718–21.

Hanotte O., Tawah C.L., Bradley D.G., Okomo M., Verjee Y., Ocheng J. & Rege J.E. (2000) Geographic distribution and frequency of a taurine *Bos taurus* and an indicine *Bos indicus* Y specific allele amongst sub-Saharan African cattle breeds. *Molecular Ecology* **9**, 387–96.

Hanotte O., Bradley D.G., Ocheng J.W., Verjee Y., Hill E.W. & Rege J.E. (2002) African pastoralism: genetic imprints of origins and migrations. *Science* **296**, 336–9.

- Kantanen J., Edwards C.J., Bradley D.G. *et al.* (2009) Maternal and paternal genealogy of Eurasian taurine cattle (*Bos taurus*). *Heredity* doi: 10.1038/hdy.2009.68.
- Kappes S.M., Keele J.W., Stone R.T., McGraw R.A., Sonstegard T.S., Smith T.P., Lopez-Corralles N.L. & Beattie C.W. (1997) A second-generation linkage map of the bovine genome. *Genome Research* 7, 235–49.
- de Knijff P. (2000) Messages through bottlenecks: on the combined use of slow and fast evolving polymorphic markers on the human Y chromosome. *American Journal of Human Genetics* 67, 1055–61.
- Li M.H., Zerabruk M., Vangen O., Olsaker I. & Kantanen J. (2007) Reduced genetic structure of north Ethiopian cattle revealed by Y-chromosome analysis. *Heredity* 98, 214–8.
- Liu W.-S., Beattie C.W. & Ponce de León F.A. (2003) Bovine Y chromosome microsatellite polymorphism. *Cytogenetic and Genome Research* 10, 53–8.
- Mathias N., Bayes M. & Tyler-Smith C. (1994) Highly informative compound haplotypes for the human Y chromosome. *Human Molecular Genetics* 3, 115–23.
- Meadows J.R.S., Hanotte O., Drögemüller C., Calvo J., Godfrey R., Coltman D., Maddox J.F., Marzanov N., Kantanen J. & Kijas J.W. (2006) Globally dispersed Y chromosomal haplotypes in wild and domestic sheep. *Animal Genetics* 37, 444–53.
- Mezzelani A., Zhang Y., Redaelli L., Castiglioni B., Leone P., Williams J.L., Toldo S.S., Wigger G., Fries R. & Ferretti L. (1995) Chromosomal localization and molecular characterization of 53 cosmid-derived bovine microsatellites. *Mammalian Genome* 6, 629–35.
- Nei M. (1987). *Molecular Evolutionary Genetics*. Columbia University Press, New York, NY.
- Pérez-Panadell L., Royo L.J., Álvarez I., Ponce de León F.A., Fernández I., Casais R. & Goyache F. (2009). Female segregation patterns of the putative Y-chromosome specific microsatellite markers INRA124 and INRA126 do not support their use for cattle population studies. *Animal Genetics* 40, 560–4.
- Sørensen A.C., Sørensen M.K. & Berg P. (2005) Inbreeding in Danish dairy cattle breeds. *Journal of Dairy Science* 88, 1865–72.
- Svensson E. & Götherström A. (2008) Temporal fluctuations of Y-chromosomal variation in *Bos taurus*. *Biological Letters* 4, 752–4.
- Troy C.S., MacHugh D.E., Bailey J.F., Magee D.A., Loftus R.T., Cunningham P., Chamberlain A.T., Sykes B.C. & Bradley D.G. (2001) Genetic evidence for Near-Eastern origins of European cattle. *Nature* 410, 1088–91.
- Vaiman D., Inam-Ghali M., Moazami-Goudarzi K., Guerin G., Grohs C., Leveziel H. & Saidi-Mehtar N. (1994) Conservation of a syntenic group of microsatellite loci between cattle and sheep. *Mammalian Genome* 5, 310–4.
- Y-Chromosome Consortium (2002) A nomenclature system for the tree of human Y-chromosomal binary haplogroups. *Genome Research* 12, 339–48.
- Zegura S.L., Karafet T.M., Zhivotovsky L.V. & Hammer M.F. (2004) High-resolution SNPs and microsatellite haplotypes point to a single, recent entry of native american Y chromosomes into the Americas. *Molecular Biology and Evolution* 21, 164–75.
- Zerjal T., Wells R.S., Yudashcheva N., Ruzibakiev R. & Tyler-Smith C. (2002) A genetic landscape reshaped by recent events: Y-chromosomal insights into Central Asia. *American Journal of Human Genetics* 71, 466–82.

## Supporting information

Additional supporting information may be found in the online version of this article.

**Table S1** Description of microsatellites tested in this study.

**Table S2** Cattle Y-chromosome haplotypes defined by six Y-specific loci and their frequencies across breeds.

**Table S3** Between breeds genetic identity computed from six Y-specific loci. Eigenvectors corresponding to each of the three principal components computed via Principal Component Analysis are also given for each breed.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery but or not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.



## ORIGINAL ARTICLE

## Multiple paternal origins of domestic cattle revealed by Y-specific interspersed multilocus microsatellites

L Pérez-Pardal<sup>1,2</sup>, LJ Royo<sup>1</sup>, A Beja-Pereira<sup>2</sup>, S Chen<sup>2</sup>, RJC Cantet<sup>3</sup>, A Traoré<sup>4</sup>, I Curik<sup>5</sup>, J Sölkner<sup>6</sup>, R Bozzi<sup>7</sup>, I Fernández<sup>1</sup>, I Álvarez<sup>1</sup>, JP Gutiérrez<sup>8</sup>, E Gómez<sup>1</sup>, FA Ponce de León<sup>9</sup> and F Goyache<sup>1</sup>

<sup>1</sup>Área de Genética y Reproducción Animal, SERIDA, Camino de Rioseco, Gijón, Spain; <sup>2</sup>CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Universidade do Porto, Campus Agrário de Vairão, Vairão, Portugal; <sup>3</sup>Departamento de Producción Animal, Facultad de Agronomía, Universidad de Buenos Aires – CONICET, Buenos Aires, Argentina; <sup>4</sup>INERA, Ouagadougou, Burkina Faso; <sup>5</sup>Animal Science Department, University of Zagreb, Svetosimunska, Zagreb, Croatia; <sup>6</sup>Institute of Livestock Sciences, University of Natural Resources and Applied Life Sciences Vienna, Vienna, Austria; <sup>7</sup>Dipartimento Biotecnologie Agrarie, Sezione Scienze Animali, Università degli Studi di Firenze, Firenze, Italy; <sup>8</sup>Departamento de Producción Animal, Universidad Complutense de Madrid, Madrid, Spain and <sup>9</sup>Department of Animal Science, University of Minnesota, St Paul, MN, USA

In this study, we show how Y-specific interspersed multilocus microsatellites, which are loci that yield several amplified bands differing in size from the same male individual and PCR reaction, are a powerful source of information for tracing the history of cattle. Our results confirm the existence of three main groups of sires, which are separated by evolutionary time and clearly predate domestication. These three groups are consistent with the haplogroups previously identified by Götherström *et al.* (2005) using five Y-specific segregating sites: Y1 and Y2 in taurine (*Bos taurus*) cattle and Y3 in zebu (*Bos indicus*) cattle. The zebu cattle cluster clearly originates from a domestication process that was geographically and temporally

separated from that of taurine clusters. Our analyses further suggest that: (i) introgression of wild sire genetic material into domesticated herds may have a significant role in the formation of modern cattle, including the formation of the Y1 haplogroup; (ii) a putative domestication event in Africa probably included local Y2-like wild sires; (iii) the West African zebu cattle Y-chromosome may have partially originated from an ancient introgression of humped cattle into Africa; and (iv) the high genetic similarity among Asian zebu sires is consistent with a single domestication process.

Heredity advance online publication, 24 March 2010; doi:10.1038/hdy.2010.30

**Keywords:** bovine; Y-chromosome; interspersed multilocus microsatellites; cattle origins; domestication

## Introduction

Cattle history, inferred from mitochondrial DNA (mtDNA) markers (Bradley *et al.*, 1996; Loftus *et al.*, 1994, 1999; Troy *et al.*, 2001; Edwards *et al.*, 2007a; Achilli *et al.*, 2008), suggests that cattle (*Bos taurus* and *Bos indicus*) domestication involved at least two genetically distinct auroch (*Bos primigenius*) species in the Near East and great Indus Valley. Most European cattle resulted from the expansion of a small cattle population from the Near East after domestication, with one (T3) of the four major haplogroups (T, T1, T2 and T3) identified in the Near East predominating in the mainland continent. However, the history of European cattle may be more complex than what was previously thought. Recent studies have identified new mitochondrial haplogroups in European cattle (Achilli *et al.*, 2008), and probable introgression events of local wild aurochs into domesticated *B. taurus* cattle (Beja-Pereira *et al.*, 2006; Achilli *et al.*, 2008). African cattle mtDNA lineages are mainly assigned to the T1 haplogroup, which is very rare in the

Middle East and Anatolia. This could suggest an independent domestication event, although the hypothesis of an African demographic expansion of a few T1-like cattle domesticated in the Near East is still arguable (Beja-Pereira *et al.*, 2006; Bradley and Magee, 2006). Regarding zebu cattle, mtDNA sequences allowed the identification of two major haplogroups: I1 and I2 (Baig *et al.*, 2005; Lai *et al.*, 2006; Lei *et al.*, 2006; Achilli *et al.*, 2008; Chen *et al.*, 2010). These two mtDNA lineages are well represented in India, suggesting that zebu domestication involved at least two different wild female populations (Baig *et al.*, 2005) or, more likely, a single domestication event in the Indus Valley with a subsequent introgression process of wild (I2) females into proto-domesticated herds (Chen *et al.*, 2010).

The DNA studies regarding cattle Y-chromosomes are comparatively rare (Verkaar *et al.*, 2004) and have mainly focused on the assessment of male-mediated migration patterns and admixture between *B. taurus* and *B. indicus* (Hanotte *et al.*, 2000; Anderung *et al.*, 2007; Edwards *et al.*, 2007b) or the assessment of differences in diversity (Ginja *et al.*, 2009; Kantanen *et al.*, 2009). Recent analyses have identified five polymorphic sites on the cattle Y-chromosome (Götherström *et al.*, 2005), allowing the identification of three clusters (Y1, Y2 and Y3) in contemporary cattle, with Y1 being more frequent in

Correspondence: Dr F Goyache, Área de Genética y Reproducción Animal, SERIDA, Camino de Rioseco 1225, Gijón, Asturias 33394, Spain.  
 E-mail: fgoyache@serida.org

Received 4 June 2009; revised 16 October 2009; accepted 22 October 2009



*B. taurus* from north-western Europe, Y2 being dominant in *B. taurus* found in southern Europe and Anatolian cattle, and Y3 being exclusive to *B. indicus*. The Y2 haplotypes would represent the taurine cattle domesticated in the Fertile Crescent, whereas the Y1 haplogroup would represent European aurochs. This would indicate a greater genetic influence of the primitive European aurochs on the formation of modern cattle breeds in Europe (Beja-Pereira et al., 2006; Achilli et al., 2008). However, ancient DNA does not generally support significant crossbreeding between local European wild cattle and domesticated cattle (Edwards et al., 2007a; Bollongino et al., 2008). In addition, medieval Scandinavian samples largely belonged to haplogroup Y2 and, overall, this scenario would not be consistent with a conserved signal from the time of domestication (Svensson and Götherström, 2008). Therefore, the strong geographical differentiation between the Y haplogroups found in modern cattle breeds in Europe requires further explanation.

Cattle Y-chromosome studies are also affected by a lack of powerful sources of information. There are limited numbers of informative segregating sites and polymorphic Y-specific microsatellites (Götherström et al., 2005; Ginja et al., 2009). Liu et al. (2003) reported the existence of several interspersed multilocus microsatellites (IMMs). These IMMs are loci that yield several amplified bands differing in size from the same male individual, using a single PCR reaction performed using a single primer pair. These loci are assumed to be interspersed and a result of the repetitive nature of the Y-chromosome. The male-specific region of the Y-chromosome is a mosaic of heterochromatic sequences and three classes of euchromatic sequences: X-transposed, X-degenerate and ampliconic (Skaletsky et al., 2003). The absence of recombination with the X-chromosome causes rapid degeneration by mutation, deletion and transposon invasion that accumulate through time due to the larger number of cell divisions required to produce male gametes, and due to inefficient repair (for example, Muller's ratchet) and inefficient selection (Charlesworth et al., 2005; Graves, 2006; Bachrog and Andolfatto, 2006; Van Laere et al., 2008). The repetitive nature of the Y-chromosome has been hypothesised to be a mechanism for maintaining Y-chromosome genes in a non-recombining environment and may arise by a number of mechanisms, including sexual antagonism, genomic conflict and hemizygous exposure (Vallender and Lahn, 2004; Murphy et al., 2006). When a section is duplicated, it may result in a new locus that could generate new alleles at the locus site over time (Butler et al., 2005). In the hemizygous state, the latter will be difficult to identify and, therefore, only its presence or absence can be identified. The new allele may be present in some individuals but absent in others, allowing each locus to be treated as a polymorphic site. Therefore, the polymorphism can be assessed as the presence/absence of a given band across individuals. However, the value of IMMs in population genetics studies has not been explored yet (Luo et al., 2007). Here, we tested a set of IMMs to ascertain their usefulness for cattle phylogeny. The consistency of the information provided by the IMMs, along with the cattle history inferred from previous genetic analyses, will be discussed.

## Materials and methods

### Selection of IMMs

A total of five microsatellite loci, previously reported as IMMs by Liu et al. (2003), were tested in both taurine and zebu samples for male specificity, paternal compatibility and repeatability scoring to ascertain their performance in cattle Y-chromosome variability assessment (Pérez-Pardal et al., 2009a). The IMM behaviour was consistent across subspecies. When amplification occurred using female DNA or identical band amplification results were not obtained with different PCR conditions, the IMM was eliminated and was not considered for further analyses (Figure 1 and Supplementary Figure S2). The PCR amplification conditions are described in Table 1. The description of the pedigrees used to test the markers is given in Supplementary Figure S2. Laboratory analyses were carried out using Automatic Sequencers ABI-310 and ABI-3130 (Applied Biosystems, Foster City, CA, USA). Size characterisation of PCR products was carried out using GenMapper software (Applied Biosystems).

### Statistical analyses

Observed haplotypes were analysed as follows:

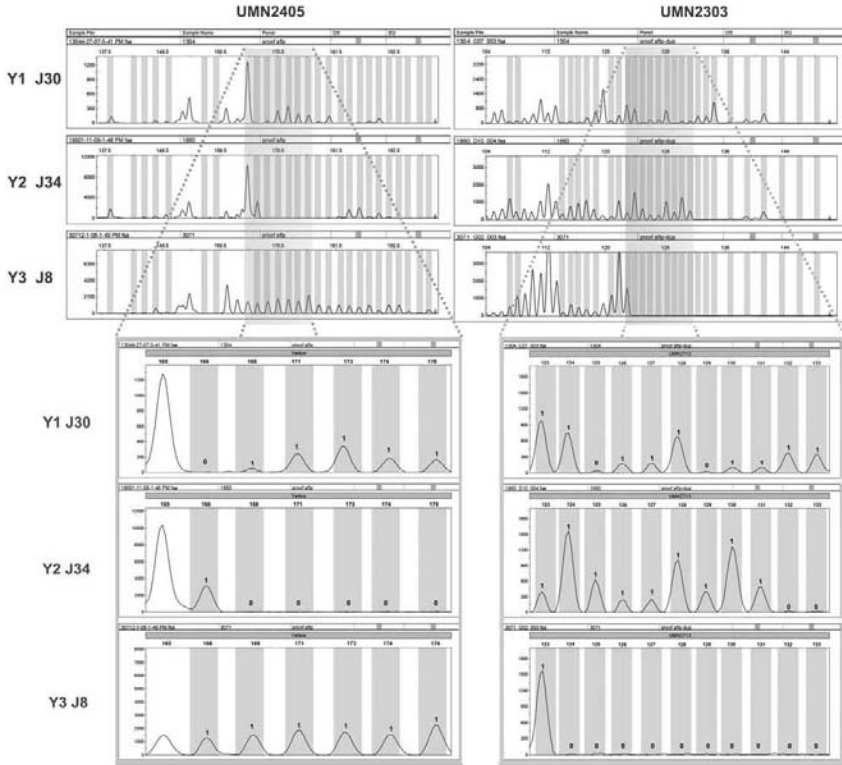
- (1) Application of the Bayesian procedures implemented in the programme MrBayes 3.1. (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003). Two MCMC runs, starting from different random trees, were completed; each with 20 000 000 generations and four chains. All sample points before reaching convergence were discarded as burn-in samples. The remaining samples were used to generate a majority rule consensus tree, in which the percentage of samples recovering any particular clade represented the clade's posterior probability (Huelsenbeck and Ronquist, 2001).
- (2) Construction of biologically meaningful phylogenies of all genotyped samples using the program Network 4.5 (available at <http://www.fluxus-engineering.com/>). To avoid reticulation, a reduced median algorithm (Bandelt et al., 1995) was used to generate a .rmf file, and the median joining network method (Bandelt et al., 1999) was applied to this file. Following the authors' recommendations (Forster et al., 2000), the 'frequency > 1' option was applied to discard singly occurring Y types. Same weights were assigned to each haplotype.
- (3) Computation of correspondence analysis using Proc Corresp of the SAS/STAT package (SAS Institute, Cary, NC, USA).

Unbiased estimates of haplotype diversity and their variances ( $V_{00}$ ) were calculated for the whole sample and for each identified haplogroup as mentioned in the study by Nei (1987) (formulas 8.5 and 8.13 therein). The s.e. value of  $h$  was calculated by taking the square root of  $V_{00}$ .

The program Arlequin 3.11 (Excoffier et al., 2005) was used to calculate the analysis of molecular variance.

Divergence times were estimated using the  $\rho$  statistic (the average number of mutations from derived haplotypes to a haplotype designated as ancestral for the haplogroup; Forster et al., 1996) using the program NETWORK 4.5. The average distance to the node of interest ( $\rho$ ) was transformed to absolute time estimates by





**Figure 1** Electropherograms of cattle Y-specific interspersed multilocus microsatellites (IMMs) UMN2405 and UMN2303. We show haplotypes J30, J34 and J8 (see Supplementary Table 2) that belong to haplogroups Y1, Y2 and Y3, respectively. Grey bars represent the loci assessed as polymorphic across all samples used. Polymorphism is assessed as the presence/absence of a given band across individuals. Examples of genotyping that were carried out are shown by zooming in on parts of the electropherograms. Band sizes are shown at the top of the zoomed-in areas. The presence or absence of a given band in a sample is assessed as 1 (at the top of the peaks) or 0.

multiplication ( $\rho \times$  years per mutation), assuming the phylogenies constructed using median-joining algorithms, as implemented in NETWORK. As a phylogeny-based statistic,  $\rho$  offers the advantage of being unbiased by demographic processes. The sampling error of  $\rho$  was approximated as  $\sqrt{\rho/n}$ , where  $n$  denotes the sample size. No published mutation rate for IMMs is available. As IMMs are defined as dominant markers, they should most probably be considered as slowly mutating markers. Therefore, we computed divergence times using the mutation rate (0.00018 per generation) reported by Forster *et al.* (2000) for slowly mutating Y-chromosome markers. The generation interval was fitted to 4.84 years. This value is the mean value of 15 different estimates reported by Gutiérrez *et al.* (2003) in eight breeds of cattle; most of them managed under very traditional conditions and not subject to sire selection programs.

The determination of ancestral haplotypes was performed by computing an internal number of square differences ( $d^2$ ), computed as the average overall IMMs typed of  $d^2 = \sum_{j=1}^n \frac{d_{ja}^2}{n}$  where  $d_{ja}^2$  is the parameter for the IMM  $A$ ,  $j_a$  and  $j_b$  are the lengths in repeat units of alleles  $a$  and  $b$  at IMM  $A$ , and  $n$  is the total number of bands at which IMM  $A$  was scored (there, 25 for IMM UMN2403 and 23 for UMN2303). Assuming a strict stepwise mutation model, the lower the value of  $d^2$ , the deeper the presence of the haplotype in the pedigree.

## Results

We tested five IMMs reported by Liu *et al.* (2003) for the following: (i) amplification only with male DNA; (ii) polymorphism; (iii) paternal compatibility; and (iv)



Table 1 Description of IMMs tested

IMM	Primer sequences [5' to 3']	No. of bands		Size range in bp	Amplification with female DNA <sup>a</sup>	Paternal compatibility	Scoring repeatability <sup>a</sup>
		Total	Polymorphic				
UMN2405 <sup>b</sup>	CCTGCCATCCATTGTGAAGA-F CTGCTTACCTGGTCAGGATT-R	30	25	138–190	No	Yes	Yes
UMN2303 <sup>b</sup>	TACTTGCTTGAGACTTACTG-F TGTTGAACACATCTGATTCTG-R	38	23	103–141	No <sup>c</sup>	Yes	Yes
UMN2713 <sup>d</sup>	GTACCTACACTAATATGTCA-F CCAAAGAAAGTTTCAGGTACA-R	25	14	92–132	No	Yes	Yes <sup>e</sup>
UMN0920 <sup>d</sup>	GTGTGAGGACTCTTGCATCTG-F CACAGGCCTAGAAGATTGAG-R	22	Not tested	266–292	No	Not tested	No
UMN1201 <sup>d</sup>	TGCTTCATCCTTCATTCCAC-F TTGTTGAGGACTCTTGCATC-R	28	Not tested	137–190	No	Not tested	No

Abbreviation: IMM, interspersed multilocus microsatellite.

Description of IMMs reported by Liu *et al.* (2003), which were tested for male specificity, paternal compatibility and scoring repeatability. Primer sequences, number of bands identified, number of polymorphic bands and corresponding base-pair size range are given. PCR was performed using a GenAmp PCR system 9700 (Applied Biosystems). The final reaction volume was 10 µl. The reaction mixture consisted of 50–100 ng of genomic DNA, 0.2 U of Taq polymerase (Biotools, Madrid, Spain), a final concentration of 0.2 µM of each primer (Sigma-Aldrich, Madrid, Spain), 200 µM of each dNTP (Biotools), and 2 mM MgCl<sub>2</sub>. The PCR protocol included an initial step of 95 °C (3 min), followed by 30 cycles of 30 s at 95 °C for DNA denaturation, 30 s at 58 °C for primer annealing, and 30 s at 72 °C for primer extension. An additional final step of 72 °C for 90 min was included to avoid the presence of the *n*-1 band when using capillary electrophoresis. Samples from the Asturiana de los Valles family were genotyped three times for all markers using different PCR amplification conditions (primer annealing at 55, 58 and 60 °C) and included two different Taq polymerase enzymes (Biotools and TaqGold; Applied Biosystems) to verify the consistency of the results.

<sup>a</sup>Amplification with female DNA and scoring repeatability were tested in both *Bos taurus* and *Bos indicus* samples.

<sup>b</sup>Total number of bands and number of polymorphic bands assessed in the 453 samples analysed in the main study.

<sup>c</sup>IMM UMN2303 amplified a 96-bp-long amplicon using female DNA, which differs in size from those amplifications obtained from male DNA. Therefore, following Liu *et al.* (2003), the Y-specificity of IMM UMN2405 was accepted.

<sup>d</sup>Total number of bands and number of polymorphic bands found in the pedigrees described above (see also Supplementary Figure S2).

<sup>e</sup>IMM UMN2713 did not show consistent scoring repeatability across cattle subspecies. Therefore, it was not used for further analyses.

correct and repeatable scoring. Two of the tested IMMs (UMN2405 and UMN2303) fulfilled these requirements and were used further in this study. The IMM UMN2405 had a total of 30 bands, 25 of which were polymorphic. Of the 38 sites, 23 were polymorphic for UMN2303 (Table 1).

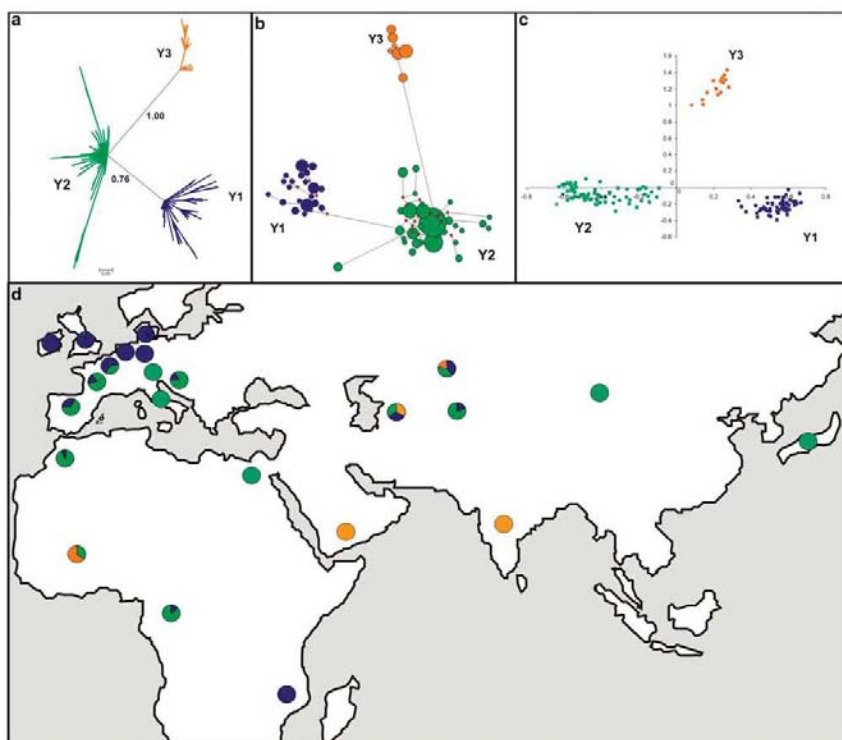
A total of 453 male individuals belonging to 52 cattle populations of Europe (351 samples), Africa (44 samples) and Asia (58 samples; Supplementary Table S1) were genotyped for the two selected IMMs. Altogether, they yielded a total of 146 haplotypes (haplotype diversity = 0.321 ± 0.005). Genotypes, haplotypes and their frequencies are provided in Supplementary Table S2.

Haplotypes were analysed by constructing evolutionary trees using Bayesian, network and correspondence analyses. All procedures allowed for the identification of three main haplogroups (Figure 2); two of these groups included only taurine samples, whereas the third was composed of only zebu samples. The consistency of these haplogroups with those previously identified by Götherström *et al.* (2005) using Y-specific single nucleotide polymorphisms was tested (see Supplementary Table S3). They were fully coincident with our haplogrouping. Therefore, the IMM-derived haplogroups will be referred to as Y1, Y2 (taurine) and Y3 (zebu). The Bayesian analysis provided a statistically high confidence value (0.76 between the taurine clusters and 1.0 among the taurine and zebu clusters). The observed haplogroups also showed high geographic consistency. The Y1 *B. taurus* haplogroup included those breeds located on the European Atlantic coastline and British

Islands, and the Y2 *B. taurus* haplogroup included those breeds spread across continental Europe, the three European (Iberian, Italian and Balkan) peninsulas and Africa. The taurine Y-chromosomes are the only cattle Y-chromosome present in Europe and are the most frequent cattle chromosomes in Africa and Northern Asia (from the Eastern Caspian Sea to Mongolia and Japan). The Y3 *B. indicus* haplogroup was the only haplogroup present in India and was well represented in Western Africa.

Haplogroup Y1 had the highest haplotype diversity (0.482 ± 0.011), whereas both the zebu (Y3) and the main taurine (Y2) haplogroups had substantially lower haplotype diversities (0.276 ± 0.006 and 0.268 ± 0.014, respectively). The Y2 haplotypes identified in African cattle are not present on other continents. Interestingly, the only Japanese Kobe and Mongolian samples available to us tended to cluster separately within haplogroup Y2 (statistical confidence: 0.98; Supplementary Figure S1) with haplotypes identified in European breeds. Most Y1 haplotypes identified in Asian and African cattle are represented in Friesian cattle, suggesting that these haplotypes did not originate on these continents.

The average number of locus differences within and between haplogroups was calculated using the program Arlequin 3.11 (Excoffier *et al.*, 2005). The higher mean pairwise differences within haplogroups for taurine were Y1 (5.08 ± 2.48) and Y2 (5.00 ± 2.44), whereas the difference was 2.58 ± 1.40 in zebu (Y3) haplotypes. The corrected mean pairwise differences between Y1 and Y2 with Y3 were 16.67 and 17.39, respectively. The difference between Y1 and Y2 was 12.06.



**Figure 2** Graphical representation of genetic variability. (a) The phylogenetic tree constructed from the 146 identified haplotypes using the Bayesian procedures implemented in the program MrBayes 3.1. (b) A network tree constructed using the program Network 4.5. (c) The two dimensions were calculated on the polymorphism of the 146 Y-chromosome haplotypes identified through correspondence analysis. The three identified haplogroups were named Y1 (in blue), Y2 (in green) and Y3 (in orange). (d) Geographic map and frequency of haplogroups (Y1: blue; Y2: green; and Y3: orange) by country (see Supplementary Table S1). The numbers of samples and haplotypes included in each haplogroup were as follows: Y1: 114/54; Y2: 281/77; and Y3: 58/15.

Analysis of molecular variance (Excoffier *et al.*, 2005) was carried out at the haplogroup, continent and geographic area levels (Table 2). Analysis of molecular variance showed that intercontinental genetic differentiation was relatively weak (explaining approximately one-third of the variability), whereas variability was greater than 50% among the defined geographical areas. Breeds accounted for less than 30% of the variation.

Divergence times were estimated using  $\rho$  procedures (Forster *et al.*, 1996) assuming three different mutation rates (see Materials and methods) and a generation interval of 4.84 years (Table 3). The estimated divergence times between haplogroups, Y3 and Y1, varied from  $131\,761.0 \pm 53\,780.0$  to  $658\,780.5 \pm 26\,889.0$  years before present (yBP). However, divergences between haplogroups Y1 and Y2 (varying from  $77\,981.0 \pm 3802.8$  to  $389\,890.5 \pm 19\,013.4$  yBP) and between Y2 and Y3 (varying

from  $75\,292.0 \pm 3802.8$  to  $376\,446.0 \pm 19\,013.4$  yBP) were substantially lower.

## Discussion

To the best of our knowledge, we have, for the first time, used Y-specific IMMs to ascertain the history of an animal species from the male side. The benefits of using IMMs in cattle are clear; when compared with the available Y-specific microsatellites and single nucleotide polymorphisms (Götherström *et al.*, 2005; Ginja *et al.*, 2009; Kantanen *et al.*, 2009; Pérez-Pardal *et al.*, 2009a, b), IMMs have higher numbers of polymorphisms and higher resolution. Very recently, using classical Y-specific microsatellites we have been able to distinguish two different Y2 haplotypic subfamilies in taurine cattle; one of them restricted to the African continent (Pérez-Pardal



**Table 2** Geographical structuring assessed using AMOVA

Type of grouping	Percentage of variance components			Total variance	$\Phi$ Statistics		
	Among groups	Among populations within groups	Within populations		$\Phi_{ST}$	$\Phi_{SC}$	$\Phi_{CT}$
Haplogroups (Y1, Y2 and Y3)	—	75.12 (2) <sup>a</sup>	24.88 (451)	9.460	0.751	—	—
Continents (Europe, Asia and Africa)	—	36.39 (2)	63.61 (451)	7.990	0.364	—	—
Geographical areas <sup>b</sup>	—	56.96 (5)	43.04 (448)	7.714	0.570	—	—
Breeds <sup>c</sup>	—	70.58 (47)	29.42 (397)	6.257	0.706	—	—
Breeds within continent	33.60 (2)	43.16 (45)	23.24 (397)	7.922	0.768	0.650	0.336
Breeds within geographic area	55.11 (5)	20.96 (42)	23.94 (397)	7.692	0.761	0.467	0.551

Abbreviation: AMOVA, analysis of molecular variance.

AMOVA subdivides the genetic diversity into hierarchical components and estimates the indices,  $\Phi$ , which are molecular equivalents of Wright's F statistics and can be interpreted as follows:  $\Phi_{ST}$ : relative divergence between populations;  $\Phi_{SC}$ : relative divergence between populations of the same group;  $\Phi_{CT}$ : relative divergence between groups. Note that the genetic variance was subdivided sequentially into one (haplogroups, continents, geographic areas or breeds) or two components (breeds within continents or geographic areas). The statistical significance of variance components and  $\Phi$  indices is evaluated by bootstrapping using 1000 replications. All estimates were statistically significant for  $P < 0.0001$ .

<sup>a</sup>Degrees of freedom are shown in parentheses.

<sup>b</sup>Geographical areas fitted were as follows (see Supplementary Table S1): Atlantic Europe (including the UK breeds, Dexter, Danish Red, Jersey, European and Holstein Friesian, Blanc-Bleu Belge, Normande, Asturian, and Tudanca cattle breeds); main European continent (including the rest of the European cattle breeds); Indian cattle (including Indian and Brahman samples); non-Indian Asia; Mediterranean Africa (including Moroccan and Egyptian samples); and Sub-Saharan Africa (including Lobi, N'Dama cattle and Landim samples).

<sup>c</sup>Populations with a sample size below three (Egypt, Landim, Mongolia and Kobe; see Supplementary Table S1) were excluded from analyses involving breeds.

**Table 3** Divergence times between ancestral haplotypes

Ancestral haplotype	J12-J19 Y3-Y2	J12-J24 Y3-Y1	J19-J24 Y2-Y1
<i>Haplogroup</i>			
P (s.d.)	14.00 (0.71)	24.50 (1.00)	14.50 (0.71)
Years (s.d.)	376 446.0 (19 013.4)	658 780.5 (26 889.0)	389 890.5 (19 013.4)

Abbreviation: IMM, interspersed multilocus microsatellite.

Divergence times between ancestral haplotypes estimated using  $\rho$  procedures as implemented in the program NETWORK 4.5, assuming that IMMs have a slow mutation rate (0.00018 per generation).

*et al.*, 2009b). However, this set of markers was not powerful enough to correctly identify and differentiate haplogroups Y1 and Y2. This result underscores the need to identify new Y-specific markers to assess genetic relationships between different bovine Y-specific haplogroups (Pérez-Pardal *et al.*, 2009b).

Moreover, the interest of this study extends beyond the bovine species, as IMMs have been identified in the Y-chromosomes of many other mammals, such as the family Felidae (Luo *et al.*, 2007) and the genus *Equus* (Wallner *et al.*, 2004). However, the use of IMMs has some constraints. First, their use is not straightforward: their selection, amplification, genotyping and scoring must be performed carefully. The second, and more important, problem is that their mutation model is currently unknown, which limits their use in evolutionary studies.

In this study, IMMs yield information about unanswered questions of cattle history from male legacies in light of previously available information: (i) what is the genetic influence of local aurochs on the formation of domesticated cattle?; (ii) do African cattle result from a local domestication event?; and (iii) are zebu cattle genetically uniform at the sire level?

Within Y2, an intriguing subcluster was identified. It includes cattle Y-chromosomes sampled in Mongolia, Japan, Northern Italy and Northern Atlantic Europe. Recently, Achilli *et al.* (2008) identified the new mtDNA haplogroup Q in Northern Italian and Korean cattle

samples, possibly resulting from the introgression of local wild aurochs. These are largely the same geographical areas in which our subcluster was found. Therefore, a likely explanation for the Y2-subcluster identified in this study is the introgression of wild sires into domesticated cattle. This introgression process would have involved poorly differentiated wild sire populations and would have happened in very distant geographical areas. This is consistent with mtDNA data that demonstrate the existence of a wild cattle population sharing many genetic features from Europe to North Asia (Achilli *et al.*, 2008). Overall, the recruitment processes of wild cattle are likely to have occurred in some parts of Europe and Northern Asia, giving genetic signatures at both mtDNA and Y-chromosome levels.

On the basis of the strong geographic signal obtained from modern samples, Götherström *et al.* (2005) proposed the possible introgression of local aurochs into European domesticated cattle. However, more recent evidence obtained from ancient and medieval samples does not support this hypothesis (Bollongino *et al.*, 2008; Svensson and Götherström, 2008), and the alternative model of a sudden introduction of domestic cattle into Central Europe without significant crossbreeding with local wild cattle remains unchallenged (Edwards *et al.*, 2007a; Bollongino *et al.*, 2008). Therefore, practical fixation of Y1 haplotypes in Northern Atlantic European cattle breeds would more likely be the result of drift



processes that occurred just before and at the moment of creation of modern cattle breeds approximately 200 yBP (Svensson and Götherström, 2008). However, this argument has an intuitive weakness because just after the formation of modern cattle breeds (approximately 100 yBP), the intense and sustained spreading of Y1-fixed cattle breeds (Friesian and British cattle) was facilitated by improved possibilities for long-distance trade and artificial insemination (approximately 50 yBP). The latter occurred all over the world, particularly across Europe. This dramatic and uncontrolled spreading process would have unavoidably erased differences among European taurine breeds at the Y-haplogroup level. It is noteworthy, that the Y1 samples identified in Africa and Northern Asia (Figure 2) are more likely to be the result of a recent introgression of European cattle rather than the expansion of a genetically heterogeneous sire population domesticated in the Fertile Crescent. Moreover, our results do not allow for the rejection of the hypothesis that introgression of local male aurochs partially influenced the formation of European cattle. This hypothesis is supported by two facts. The first is that the haplotypic diversity in the Y1 cluster is substantially higher than that observed for haplogroups Y2 and Y3, which are expected to be the result of domestication. This would suggest that the occurrence of a few local introgression events of Y1-related sires rather than a domestication process resulting in a much more intense bottleneck. The second fact is that ancient cattle management differences between northern and southern European cattle holders are supported by observed differences between northern and southern European cattle populations at the mtDNA level (Beja-Pereira et al., 2006), and evidence of co-evolution of cattle milk protein genes and human lactase in Northern Europe (Beja-Pereira et al., 2003). Our results support the idea that cattle management in Atlantic Europe might include the use of wild males with domesticated cows.

The question of the origins of both African taurine and zebu cattle has been the subject of intense debate and remains controversial. Our study confirms the existence of a Y2 haplotypic subfamily in African cattle (Pérez-Pardal et al., 2009b). Consistent with some archaeological findings (Wendorf and Schild, 1994), Bradley and Magee (2006) claim that African taurine cattle are not a simple subset of those domesticated in the Near East, due to the following three facts: (i) the mtDNA composition of cattle from these two areas is qualitatively different, with African samples belonging to the T1 haplogroup and showing a large proportion of unique haplotypes (Loftus et al., 1994; Bradley et al., 1996; Troy et al., 2001); (ii) there is a significant component of microsatellite variation within African cattle that seems to be indigenous in origin (Hanotte et al., 2002); and (iii) microsatellite variation also suggests that gene flow from the Near East into Northeast Africa was predominantly male-driven (MacHugh et al., 1997). Our results support this hypothesis because: (i) most haplotypes identified in African taurines are assigned to haplogroup Y2 and are not present in other continents; and (ii) the African Y2 samples show high haplotypic diversity ( $0.708 \pm 0.069$ ). At the mtDNA level, African cattle do not present the lack of variation associated with a severe foundation bottleneck, which would be necessary to fix a haplotype (T1) family after the domestication event occurred in the

Fertile Crescent (Bradley and Magee, 2006). This evidence can also be applied to our findings regarding the Y-chromosome.

The origin of African cattle is also linked to humped zebu cattle. Archaeological evidence suggests introgression of the humped zebu-like cattle into Africa approximately 3000 yBP (Payne, 1970). However, a major wave of *B. indicus* introgression in Africa may have started around the end of the seventh century AD in the Arab settlements along the East coast of Africa (Epstein, 1971), with subsequent spreading through zebu male-mediated interbreeding with African taurine females (Hanotte et al., 2000, 2002). Four of the six Y3 haplotypes identified in West African cattle tend to cluster together and are not present in Asian zebu sires (Supplementary Figure S1). This finding indicates a non-recent introgression of zebu cattle into Africa. Indeed, West African cattle share haplotypes with Asian zebu samples. However, West African cattle are often regarded as modern descendants of the founding African *B. taurus* cattle, which have been less affected by recent *B. indicus* introgression into the continent (Hanotte et al., 2002; Edwards et al., 2007b). Therefore, West African sanga cattle (the technical designation for the cross between taurine and zebu cattle) could be representative of a first ancient introgression of zebu cattle into Africa.

The scenario seen in our Asian zebu samples is quite different from that of cattle on the other continents. Most Y3 haplotypes observed in Asian zebu sires are present in the Indian populations, which leads to low haplotypic diversity ( $0.256 \pm 0.017$ ). At the maternal genetic level, a major component of zebu cattle diversity is retained by two distinct mitochondrial haplogroups (I1 and I2), suggesting two different domestication events or a single domestication process involving two different wild female populations (Baig et al., 2005; Lai et al., 2006; Chen et al., 2010). In this study, the sampling size for Asian zebu sires was limited. However, Indian samples were obtained in areas in which both I1 and I2 mtDNA clades are well represented (Chen et al., 2010). Despite this, haplotypic diversity is extremely low, suggesting high genetic homogeneity of Asian male zebus. This would support the hypothesis of a single domestication event for *B. indicus*. Recently, Chen et al. (2010) suggested that I1-like zebu cattle were domesticated in a single event in the Indus Valley (approximately 6000 yBP), with subsequent recruitment of mitochondrial I2 wild zebu females approximately 3500 yBP. Our results suggest that this introgression event of zebu wild females into proto-domesticated herds would have occurred with previously domesticated zebu sires.

To the best of our knowledge, our IMM data, for the first time, allowed the estimation of divergence times between cattle subspecies (*B. taurus* and *B. indicus*) on the male genetic path. Recent analyses of the complete *B. taurus* and *B. indicus* mitochondrial genome sequences estimated their divergence times to be 1.7–2.0 million years (Hiendleder et al., 2008). However, using a fragment of the HVI region of mtDNA, most divergence times between the two subspecies were estimated to be 200 000–1 000 000 yBP (Loftus et al., 1994 and others). This mtDNA fragment gives information that is comparable in amount to that given by our IMM. Current estimates are highly consistent with the generally accepted range of divergence times (Loftus et al., 1994). However,



divergence times estimated between the Y1–Y2 and Y2–Y3 haplogroups are consistently lower than the estimated time between Y1–Y3. Overall, this would indicate a recent fragmentation of the wild populations that predated domestication, and a lack of differentiation between geographically distant wild sire populations.

In summary, we have demonstrated the usefulness of Y-specific IMMs to complement and add information to mtDNA-based studies regarding the history of the species. The divergent sire lineages, characterised here by Y-haplogroups, align with geographic areas and support previously suggested origins of cattle and directions of spread (Bradley *et al.*, 1996; Troy *et al.*, 2001; Baig *et al.*, 2005; Lai *et al.*, 2006; Edwards *et al.*, 2007a; Achilli *et al.*, 2008). In light of the above evidence, the most parsimonious interpretation of our data is that cattle have multiple paternal origins, possibly arising through multiple independent domestications. However, it is also possible that many of the paternal lineages identified in this study originated through introgression and not through separate domestication events.

### Conflict of interest

The authors declare no conflict of interest.

### Acknowledgements

We thank CC Fioretti (Cabaña Las Lilas, Buenos Aires, Argentina), E Miura (Unión de Criadores de Toros de Lidia, Madrid), PJ Azor (Universidad de Córdoba, Spain), J Arranz (Universidad de León, Spain), D Pirotin (Université de Liège), D de Petris (Università degli Studi di Firenze, Italy), T Lecomte (JVL Company, Belgium-Congo), B Scherf (FAO, Rome, Italy), J Jordana (Universitat Autònoma de Barcelona), and L Domínguez (Universidad Complutense de Madrid, Spain) for their support and help with sampling. This study was partially supported by grants from MICIN no. CGL2005-03761/BOS and no. CGL2008-03949/BOS, from the Government of Principado de Asturias no. IB09-114 and FCT grant POCI/CVT/56758/2004. LP-P is supported by grant MICINN BES-2006-13545. ABP and SC are supported by FCT grants SFRH/BPD/38096/2007 and SFRH/BPD/26802/2006, respectively. APdL is supported by USDA-Hatch project MIN-16-019.

### References

Achilli A, Olivieri A, Pellecchia M, Uboldi C, Colli L, Al-Zahery N *et al.* (2008). Mitochondrial genomes of extinct aurochs survive in domestic cattle. *Curr Biol* 18: R157–R158.  
 Anderung C, Hellborg L, Seddon J, Hanotte O, Götherström A (2007). Investigation of X- and Y-specific single nucleotide polymorphisms in taurine (*Bos taurus*) and indicine (*Bos indicus*) cattle. *Anim Genet* 38: 595–600.  
 Bachrog D, Andolfatto P (2006). Selection, recombination and demographic history in *Drosophila miranda*. *Genetics* 174: 2045–2059.  
 Baig M, Beja-Pereira A, Mohammad R, Kulkarni K, Farah S, Luikart G (2005). Phylogeography and origin of Indian domestic cattle. *Curr Sci* 89: 38–40.  
 Bandelt HJ, Forster P, Röhl A (1999). Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* 16: 37–48.

Bandelt H-J, Forster P, Sykes BC, Richards MB (1995). Mitochondrial portraits of human populations. *Genetics* 141: 743–753.  
 Bradley DG, MacHugh DE, Cunningham P, Loftus RT (1996). Mitochondrial diversity and the origins of African and European cattle. *Proc Natl Acad Sci USA* 93: 5131–5135.  
 Bradley DG, Magee DA (2006). Genetics and origins of domestic cattle. In: Zeder MA, Bradley DG, Emshwiler E, Smith BD (eds). *Documenting domestication New genetic and archaeological paradigms*. University of California Press: Berkeley-Los Angeles-London, pp 317–328.  
 Beja-Pereira A, Caramelli D, Lalueza C, Vernesi C, Ferrand N, Sampietro L *et al.* (2006). Genetic evidence for multiple origins of European cattle in Near-East, Africa, and Europe. *Proc Natl Acad Sci USA* 103: 8113–8118.  
 Beja-Pereira A, Luikart G, England PR, Bradley DG, Jann OC, Bertorelle G *et al.* (2003). Gene-culture coevolution between cattle milk protein genes and human lactase genes. *Nat Genet* 35: 311–313.  
 Bollongino R, Elsner J, Vigne J-D, Burger J (2008). Y-SNPs do not indicate hybridisation between European aurochs and domestic cattle. *PLoS ONE* 3, <http://dx.doi.org/10.1371/journal.pone.0003418>.  
 Butler JM, Decker AE, Kline MC, Vallone PM (2005). Chromosomal duplications along the Y-chromosome and their potential impact on Y-STR interpretation. *J Forensic Sci* 50: 853–859.  
 Charlesworth D, Charlesworth B, Marais G (2005). Steps in the evolution of heteromorphic sex chromosomes. *Heredity* 95: 118–128.  
 Chen S, Lin B-Z, Baig M, Mitra B, Lopes RJ, Tarroso P *et al.* (2010). Zebu cattle are an exclusive legacy of the South Asia Neolithic. *Mol Biol Evol* 27: 1–6.  
 Edwards CJ, Baird JE, MacHugh DE (2007b). Taurine and zebu admixture in Near Eastern cattle: a comparison of mitochondrial, autosomal and Y-chromosomal data. *Anim Genet* 38: 520–524.  
 Edwards CJ, Bollongino R, Scheu A, Chamberlain A, Tresselt A, Vigne J-D *et al.* (2007a). Mitochondrial DNA analysis shows a near Eastern Neolithic origin for domestic cattle and no indication of domestication of European aurochs. *Proc Biol Sci* 274: 1377–1385.  
 Epstein H (1971). *The Origins of the Domestic Animals of Africa*, vol. 1. African Publishing Corporation: London.  
 Excoffier L, Laval G, Schneider S (2005). Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evol Bioinf Online* 1: 47–50.  
 Forster P, Harding R, Torroni A, Bandelt H-J (1996). Origin and evolution of native American mtDNA variation: a reappraisal. *Am J Human Genet* 59: 935–945.  
 Forster P, Röhl A, Lünemann P, Brinkmann C, Zerjal T, Tyler-Smith C *et al.* (2000). A short tandem repeat-based phylogeny for the human Y chromosome. *Am J Hum Genet* 67: 182–196.  
 Ginja C, Telo da Gama L, Penedo MCT (2009). Y Chromosome haplotype analysis in Portuguese cattle breeds using SNPs and STRs. *J Hered* 100: 148–157.  
 Götherström A, Anderung C, Hellborg L, Elburg R, Smith C, Bradley DG *et al.* (2005). Cattle domestication in the near East was followed by hybridization with aurochs bulls in Europe. *Proc R Soc B* 272: 2345–2350.  
 Graves JAM (2006). Sex chromosome specialization and degeneration in mammals. *Cell* 124: 901–914.  
 Gutiérrez JP, Altarriba J, Díaz C, Quintanilla R, Cañón J, Piedrafita J (2003). Pedigree analysis of eight Spanish beef cattle breeds. *Genet Sel Evol* 35: 43–63.  
 Hanotte O, Bradley DG, Ochieng JW, Verjee Y, Hill EW, Rege JE (2002). African pastoralism: genetic imprints of origins and migrations. *Science* 296: 336–339.  
 Hanotte O, Tawah CL, Bradley DG, Okomo M, Verjee Y, Ochieng J *et al.* (2000). Geographic distribution and frequency of a taurine *Bos taurus* and an indicine *Bos indicus* Y-specific



- allele amongst sub-Saharan African cattle breeds. *Mol Ecol* 9: 387–396.
- Hiendleder S, Lewalski H, Janke A (2008). Complete mitochondrial genomes of *Bos taurus* and *Bos indicus* provide new insights into intra-species variation, taxonomy and domestication. *Cytogenet Genome Res* 120: 150–157.
- Huelsenbeck JP, Ronquist F (2001). MRBAYES: Bayesian inference of phylogeny. *Bioinformatics* 17: 754–755.
- Kantanen J, Edwards CJ, Bradley DG, Vinalass H, Thessler S, Ivanova Z et al. (2009). Maternal and paternal genealogy of Eurasian taurine cattle (*Bos taurus*). *Heredity* 103: 404–415.
- Lai SJ, Liu YP, Liu YX, Li XW, Yao YG (2006). Genetic diversity and origin of Chinese cattle revealed by mtDNA D-loop sequence variation. *Mol Phylog Evol* 38: 146–154.
- Lei CZ, Chen H, Zhang HC, Cai X, Liu RY, Luo LY et al. (2006). Origin and phylogeographical structure of Chinese cattle. *Anim Genet* 37: 579–582.
- Liu W-S, Beattie CW, Ponce de León FA (2003). Bovine Y chromosome microsatellite polymorphism. *Cytogenet Genome Res* 10: 53–58.
- Loftus RT, Ertugrul O, Harba AH, El-Barody MAA, MacHugh DE, Park SDE et al. (1999). A microsatellite survey of cattle from a centre of origin: the Near East. *Mol Ecol* 8: 2015–2022.
- Loftus RT, MacHugh DE, Bradley DG, Sharp PM, Cunningham P (1994). Evidence for two independent domestications of cattle. *Proc Natl Acad Sci USA* 91: 2757–2761.
- Luo S-J, Johnson WE, David VA, Menotti-Raymond M, Stanyon R, Cai QX et al. (2007). Development of Y chromosome intra-specific polymorphic markers in the Felidae. *J Hered* 98: 400–413.
- MacHugh DE, Shriver MD, Loftus RT, Cunningham P, Bradley DG (1997). Microsatellite DNA variation and the evolution, domestications and phylogeography of taurine and zebu cattle. *Genetics* 146: 1071–1086.
- Murphy WJ, Wilkerson AJ, Raudsepp T, Agarwala R, Schaffer AA, Stanyon R et al. (2006). Novel gene acquisition on carnivore Y chromosomes. *PLoS Genet* 2: 353–363.
- Nei M (1987). *Molecular Evolutionary Genetics*. Columbia University Press: New York.
- Payne WJA (1970). *Cattle Production in the Tropics*, 1st edn. Longman: London.
- Pérez-Pardal L, Royo LJ, Ivarez I, Ponce de León FA, Fernández I, Casais R et al. (2009a). Female segregation patterns of the putative Y-chromosome specific microsatellite markers INRA124 and INRA126 do not support their use for cattle population studies. *Anim Genet* 40: 560–564.
- Pérez-Pardal L, Royo LJ, Beja-Pereira A, Curik I, Traoré A, Fernández I et al. (2009b). Y-specific microsatellites reveal an African subfamily in taurine (*Bos taurus*) cattle. *Anim Genet*, doi: 10.1111/j.1365-2052.2009.01982.x.
- Ronquist F, Huelsenbeck JP (2003). MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572–1574.
- Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, Cordum HS, Hillier L, Brown LG et al. (2003). The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature* 423: 825–837.
- Svensson E, Götherström A (2008). Temporal fluctuations of Y-chromosomal variation in *Bos taurus*. *Biol Lett* 4: 752–754.
- Troy CS, MacHugh DE, Bailey JE, Magee DA, Loftus RT, Cunningham P et al. (2001). Genetic evidence for near-Eastern origins of European cattle. *Nature* 410: 1088–1091.
- Vallender EJ, Lahn BT (2004). How mammalian sex chromosomes acquired their peculiar gene content. *Bioessays* 26: 159–169.
- Van Laere AS, Coppieters W, Georges M (2008). Characterization of the bovine pseudoautosomal boundary: documenting the evolutionary history of mammalian sex chromosomes. *Genome Res* 18: 1884–1895.
- Verkaar ELC, Nijman IJ, Beeke M, Hanekamp E, Lenstra JA (2004). Maternal and paternal lineages in cross-breeding bovine species. Has wisent a hybrid origin? *Mol Biol Evol* 21: 1165–1170.
- Wallner B, Piumi F, Brem G, Müller B, Achmann R (2004). Isolation of Y chromosome-specific microsatellites in the horse and cross-species amplification in the genus *Equus*. *J Hered* 95: 158–164.
- Wendorf F, Schild R (1994). Are the early Holocene cattle in the Eastern Saharan domestic or wild? *Ecol Anthropol* 3: 118–128.

Supplementary Information accompanies the paper on Heredity website (<http://www.nature.com/hdy>)



## Genetic structure of the bovine Y- specific microsatellite UMN0103 reflects the genetic history of the species

L. Pérez-Pardal 1, C. Ginja 2,6, L. J. Royo 1, I. Álvarez 1, I. Fernández 1, A del Valle 2, A. Traore 3; A. Ponce de Leon 4, A Beja-Pereira 5, M. C. T Penedo 2, F. Goyache 1.

- 1 Área de Genética y Reproducción Animal, SERIDA, Camino de Rioseco 1225 33394 Gijón, Spain.
- 2 Veterinary Genetics Laboratory, University of California, One Shields Avenue, Davis, CA 95616, USA.
- 3 INERA, 04 BP 8645 Ouagadougou 04, Burkina Faso.
- 4 Department of Animal Science, University of Minnesota, St. Paul, MN, 55108, USA
- 5 CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Universidade do Porto, Campus Agrário de Vairao, 4485-661 Vairao, Portugal
- 6 Departamento de Genética, Melhoramento Animal e Reprodução, Instituto Nacional dos Recursos Biológicos, Fonte Boa, 2005-048 Vale de Santarém, Portugal.

### ABSTRACT

The structure of the Y-specific microsatellite UMN0103 was analysed on 35 *Bos taurus* and *Bos indicus* sire samples, belonging to haplogroups Y1, Y2 and Y3, from 16 European, African and Asian cattle populations. The aim was to ascertain the number of loci and their variation related to the phylogeographic history of the different cattle strains. Two PCR protocols refer as a standard and hot-start yield two different amplification patterns for UMN0103. Analyses using standard PCR produced 2 different amplified bands in all samples except for most African Y2 sires which showed only one. The stringent hot-start PCR amplified two bands in Y3 samples but only one band in Y1 and Y2 sires. The constant amplified bands across samples was the larger band observed in European Y2 sires and the smaller allele in Y1 samples. This would suggest the existence of between-loci differences in the annealing primer site. Results point out that mutations in evolutionary duplications of the marker UMN0103 have occurred in cattle Y-chromosomes (Y1, Y2 and Y3) and adds more support to the genetic uniqueness of an African-specific subfamily within the cattle Y2 haplogroup. A second sequence difference was identified in Y3 sires which showed a point mutation A to G in the larger band [(CA)<sub>3</sub>CG (CA)<sub>n</sub>].

**Keywords** cattle, UMN0103, Interspersed multilocus microsatellites, Y-chromosome, Y-specific markers.

**Main text:**

Y-chromosome microsatellites have proven to be useful in a number of fields including phylogenetic, genealogic and forensic studies (Gusmao *et al.* 2005; Bianchi *et al.* 1998; Kayser *et al.* 1997). So far, a small number of Y-specific microsatellites have been identified in cattle. Most of them, however, are in the PAR region, are monomorphic or occur in several copies (Liu & Ponce de León, 2007), making its use difficult to infer some population genetics parameters (Pérez-Pardal *et al.* 2009). The presence of several copies in some markers has been frequently reported in humans (Schneider 1998). In cattle, the Y-specific microsatellite UMN0103 (Liu *et al.* 2003) has been used in diversity studies (Ginja *et al.* 2009; Perez-Pardal *et al.* 2010a), and reported to present two copies both in *B. taurus* and *B. indicus* and genotyped as two independent *loci* (Perez-Pardal *et al.* 2010a). However, other authors reported some differences: Ginja *et al.* (2008) found one copy in *B. taurus* and two copies in *B. indicus* and Cai *et al.* (2006) reported multiple copies on this marker. These discrepancies forced us to: i) ascertain the genetic structure of the marker; and ii) assess its information content to evaluate the genetic history of cattle.

A total of 35 sire samples belonging to individuals from Europe, Asia and Africa were genotyped for the Y-specific SNPs described by Götherstrom *et al.* (2005) following Pérez-Pardal *et al.* (2010a) to assign them to the haplogroups Y1, Y2 (belonging to *B. taurus*) and Y3 (*B. indicus*). Marker UMN0103 was amplified following the standard PCR conditions described in Pérez-Pardal *et al.* (2010a; Table 1), and also using the stringent hot-start PCR conditions described in Ginja *et al.* (2008; Table 1).

Two PCR protocols used at laboratories reporting differences in amplification products for UMN0103 were compared to ascertain their reproducibility of amplification patterns and possibly explain observed differences. Analyses based on an standard PCR protocol produced 2 bands differing in size (the issue here is that these two bands per sample are different *loci* across all samples, meaning there are two bands but they are not necessarily the same bands) in each sample except for 15 out of the 16 African Y2 sires which had only one band (see Figure S1). PCR products were cloned and sequenced. All the sequenced bands showed the same sequence variation (CA)<sub>n</sub> except for the largest band sequenced in the Y3 sires that showed a difference in the repeat motif (CA)<sub>3</sub> CG (CA)<sub>n</sub>. Except for an isolated point mutation on the more sized allele of *B. indicus* sires, no systematic differences could be observed either in the repeat or in the flanking regions between the two fragments in a given animal. Therefore, only in *B. indicus* it was possible to discriminate between alleles belonging to two different *loci* of UMN0103.

BLAST was used to compare our sequences with those in GenBank (Table S1), where only *B. taurus* sequences were available. Most clones had two similar

regions, one of them showing 99% identity with our sequence and the other showing 88 % identity. Only one clone (acc n° AC172950.4) showed 100% identity with the annealing primers zone. Two clones (acc n° AC226431.3; acc n° AC167351.3) showed 100% identity with the Reverse primer annealing sequence but present an insertion (AT) in the Forward primer annealing sequence. Analyses carried out using the program RepeatMasker revealed that the CAn surroundings included SINE, LINE, LTR and other repetitive sequences.

PCR products obtained using the stringent hot-start protocol gave two different bands in Y3 samples but only one band in Y1 and Y2 sires, coinciding in all cases with the larger allele in European Y2 sires and the smaller allele in Y1 animals.

The analyses carried out clearly demonstrate that the marker UMN0103 has, in most cattle populations, two different *loci* giving information that can be used for diversity. Because of the homogeneity of the studied sequences the two *loci* identified in the marker could not be isolated. However, indirect evidence supports the statement above. The existence of two different copies of the marker on Y3 samples was demonstrated by the different sequence of the larger allele. PCR artefacts (the new test will determine this) can be excluded as a cause because this sequence difference has been confirmed by sequencing several independent clones from each individual. Moreover, the disappearance of one band in Y1 and European Y2 samples when stringent PCR conditions are applied, suggests the existence of between-loci differences at the annealing primer site, possibly an insertion as suggested by clones AC226431.3 and AC167351.3. Moreover, Pérez-Pardal *et al.* (2010a) genotyping 216 Y1 and 365 Y2 sires from European cattle breeds for the marker UMN0103 found strong differences in allelic frequencies (see Table S2 of that paper): the most frequent alleles in Y1 samples were 122 (44.2%) and 126 (45.4%) while in Y2 samples the most frequent alleles were 120 (47.8%) and 130 (48.5%). Only one copy was detected in Y2 African animals, with no differences in sequence. This means that these animals only have one copy or have multiple copies with identical microsatellite length that would escape detection.

In *B. indicus* we have identified a within-locus sequence variation, which is has not benn identified in *B. taurus*, thus underscoring differences among major cattle groups. Moreover, when stringent PCR protocols are used the lager allele can not be identified in Y1 samples while European Y2 samples follow the opposite pattern. It would mean that the evolution of these two taurine Y-chromosomes was different at this marker, possibly fixing mutations in different annealing primer sites. The case of the African Y2 samples is interesting enough because, except for one sample, no differences in within-marker allele sizes could be identified.

This scenario completely agrees with the general scenario described by Göthertrom *et al.* (2005) as recently revisited by Pérez-Pardal *et al.* (2010a,b) via adding to the SNPs information reported by the former authors the genetic variation of cattle Y-specific microsatellites and Interspersed Multilocus Microsatellites (IMM). To the three Y-haplogroups described by Göthertrom *et al.* (2005) using SNPs the addition of microsatellite and IMM markers allows identifying an African-specific subfamily within the Y2 haplogroup.

Additionally, the results obtained contribute to increase our understanding of IMMs (loci yielding several amplified bands differing in size, from the same male individual, using a single PCR reaction performed with a single primer pair). Although IMMs may have 25 different copies or more (Pérez-Pardal *et al.* 2010b) one can hypothesise that UMN0103 is an IMM in an early moment of evolution. The existence of markers amplifying several copies is closely linked with the repetitive nature of the Y-chromosome that has been hypothesized to be a mechanism for maintaining Y-specific genes in a non-recombining environment (Vallender & Lahn 2004; Murphy *et al.* 2006). Surrounds of sequences that show identity with our sequence are full of mobile elements (MEs) as SINEs, LINEs and LTR. Different hypotheses may explain the multiplication events. The two main pathways are by means of transposition of MEs and by recombination. The question remains as to whether we are dealing with two separate processes, that both lead to redundancy, or if it is an integrated mechanism (Van't Hof *et al.* 2007). In our case the presence of MEs surrounding our sequence target could explain the presence of two copies in this marker.

## Acknowledgements

This work was partially supported by projects MCINN CGL2008-03949/BOS and PCTI-Asturias 2006-2009 IB09-114. LP-P is supported by grant MICINN BES-2006-13545. APdeL is supported by the USDA-Hatch project MIN-16-019.

## References

- Bianchi N.O., Catanesi C.I., Baillet G., *et al.* (1998) Characterization of ancestral and derived Y-chromosome haplotypes of New World native populations, *American Journal of Human Genetics* **63**, 1862–71.
- Cai X., Chen H., Wang S., Xue K. & Lei C. (2006) Polymorphisms of two Y chromosome microsatellites in Chinese cattle. *Genetics Selection Evolution* **38**, 525–34.



Ginja C., Telo da Gama L. & Penedo M.C.T. (2009) Y Chromosome haplotype analysis in Portuguese cattle breeds using SNPs and STRs. *Journal of Heredity* **100**, 148–57.

Götherström A., Anderung C., Hellborg L., Elburg R., Smith C., Bradley D.G. & Ellegren H. (2005) Cattle domestication in the near East was followed by hybridization with aurochs bulls in Europe. *Proceedings of the Royal Society B* **272**, 2345–50.

Gusmao L., Sanchez-Diz P., Calafell F., *et al.*, (2005) Mutation rates at Y chromosome specific microsatellites. *Human Mutation* **26**, 520–8.

M. Kayser, A. Caglia, D. Corach, *et al.*, (1997) Evaluation of Y-chromosomal STRs: a multicenter study. *International Journal of Legal Medicine* **110**, 125–33.

Liu W.-S., Beattie C.W. & Ponce de León F.A. (2003) Bovine Y chromosome microsatellite polymorphism. *Cytogenetic and Genome Research* **10**, 53–8.

Liu & Ponce de León F.A. (2007) Mapping of the Bovine Y Chromosome. *Electronic Journal of Biology* **1**, 5-12.

Murphy W.J., Wilkerson A.J.P., Raudsepp T., *et al.* (2006) Novel gene acquisition on carnivore Y chromosomes. *PLoS Genetics* **2**, 353–63.

Pérez-Pardal L., Royo L.J., Álvarez I., Ponce de León F.A., Fernández I., Casais R. & Goyache F. (2009). Female segregation patterns of the putative Y-chromosome specific microsatellite markers INRA124 and INRA126 do not support their use for cattle population studies. *Animal Genetics* **40**, 560–4.

Pérez-Pardal, L., Royo, L.J., Beja-Pereira, A., *et al.* (2010a) Y-specific microsatellites reveal an African subfamily in taurine (*Bos taurus*) cattle. *Animal Genetics*, **41**, 232–41.

Pérez-Pardal L., Royo L.J., Beja-Pereira A., *et al.* (2010b) Multiple paternal origins of domestic cattle revealed by Y-specific interspersed multilocus microsatellites. *Heredity* doi:10.1038/hdy.2010.30

Schneider, P.M. (1998) Tandem repeat structure of the duplicated Y chromosomal STR locus DYS385 and frequency studies in the German and three Asian populations. *Forensic Sci Intl.* **97**, 61–70

Vallender E.J. & Lahn B.T. (2004) How mammalian sex chromosomes acquired their peculiar gene content. *Bioessays* **26**, 159–69.

Van't Hof A.E., Brakefield P.M., Saccheri I.J. & Zwaan B.J. (2007) Evolutionary dynamics of multilocus microsatellite arrangements in the genome of the butterfly *Bicyclus anynana*, with implications for other Lepidoptera. *Heredity* 98, 320–328.

### **Supporting information**

Additional supporting information may be found in the online version of this article.

**Table S1** Clone AC167351.3 (position 3901 to 3946) analyzed using the RepeatMasker program.

**Figure S1** Genmapper image of UMN0103 amplification in 4 Animals belonging haplogroups Y1,Y2 and Y3.

**Table 1** Description of the haplotype analyses done per Götherström *et al.*'s (2005), geographical sampling area, breed and PCR conditions. Primer sequences were obtained from Liu *et al.* (2003). The number of samples per genotype and continent (N) is also given. PCR fragments obtained using both PCR protocols were separated by capillary electrophoresis on ABI 3730 instruments (Applied Biosystems, Barcelona, Spain). Allele sizes were determined using Genemapper software (Applied Biosystems, Barcelona, Spain). Allele sizes were standardized to the Pérez-Pardal *et al.*'s (2010a) nomenclature. Allele sizes reported by Ginja *et al.* (2008) were 2 bp bigger than those reported here. Products of standard PCR were cloned with the Topo TA cloning Kit (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. For each animal and amplified band a minimum of 3 different clones were sequenced with the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA) on an ABI 3730 instrument (Applied Biosystems, Barcelona, Spain). Sequences were analyzed using the Seqman II software v6.1 (DNASTAR Inc., Madison, WI).

Haplogroup	Sampling area	Breed	N	PCR conditions		
				Standard <sup>1</sup>	Stringent	
Y1	Europe		5	112/126	112	
		Belgium	Blanc Bleu Belge	1	112/126	112
		Spain	Asturiana de la Montaña	1	112/126	112
			Tudanco	1	112/126	112
		The Netherdlands	Holstein	1	112/126	112
		UK	Hereford	1	112/126	112
Y2	Europe		5			
		Croatia	Simental	1	122/130	130
		France	Charolaise	1	122/130	130
		Spain	Asturiana de los valles	1	120/130	130
			Morucha	1	114/130	130
		UK	Jersey	1	126/132	132
	Africa		16			
		Burkina Faso	Lobi	2	126	126
				5	128	128
				2	130	130
		Congo	N'Dama	5	122	122
				1	122/128	- <sup>2</sup>
				1	128	128
Y3	Africa		6			
		Burkina Faso	Lobi	2	114/122g	114/122
			Peul	3	112/124g	112/124
			Gudali	1	114/122g	114/122
	Asia		2			
	India		Brahman	1	114/120g	114/120
			NorthEast India	1	112/122g	112/122

<sup>1</sup>g-allele presenting the following formula: (CA)<sub>3</sub> CG (CA)<sub>n</sub>

<sup>2</sup>failure in amplification





*Discusión*



#### 4.1 Contextualización

El estudio del cromosoma Y presenta grandes dificultades debido a características específicas como la presencia de secuencias repetitivas y la falta de recombinación con el cromosoma X. Pese a ello, el análisis del cromosoma Y sigue siendo la fuente de información que permite conocer la historia de las líneas paternas en las especies mamíferas. La especie en que se han desarrollado un mayor número de trabajos sobre la historia del cromosoma Y es la especie humana (Zerjal *et al.* 2002), pero ni siquiera en esta especie se conoce la secuencia del cromosoma Y en su totalidad (Skalestky *et al.*, 2003). En ganado bovino, la secuencia del cromosoma Y no se ha publicado junto con la del resto de su genoma (Bovine Hapmap Consortium, 2009; Tellman *et al.*, 2009), aunque se siguen desarrollando trabajos para conocer el contenido génico de este cromosoma (Vogel *et al.*, 1997; Kato *et al.*, 1995; Poloumienko 2004; Hellborg *et al.*, 2003; Van Laere *et al.*, 2008), que no ha sido determinado.

El objetivo general de este trabajo era obtener una batería de marcadores que nos permitiese establecer relaciones filogenéticas por la vía masculina entre bovinos domésticos. Los trabajos publicados hasta entonces utilizaban un pequeño número de marcadores de tipo microsatélite, cuya principal utilidad consistía en diferenciar los cromosomas Y de origen *B. indicus* y *B. taurus* (Li *et al.*, 2007; Hanotte *et al.*, 2000; Edwards *et al.* 2000, 2007; Giovambattista *et al.* 2000). Posteriormente se publicaron un conjunto de cinco SNP que permiten distinguir tres linajes diferentes en el cromosoma Y bovino (Götherström *et al.*, 2005) y en 2008 se combinó por primera vez, una colección de marcadores microsatélites con el uso de los SNP identificados por Götherström *et al.*, (2005) aunque en una pequeña colección de muestras (Ginja *et al.*, 2008).

El proyecto que sirvió de base inicial a esta Memoria planteaba la utilización de microsatélites ya conocidos (Bishop *et al.*, 1994; Vaiman *et al.*, 1994;

Mezzelani *et al.*, 1995; Kappes *et al.*, 1996; Liu *et al.*, 2003) para realizar filogenia del cromosoma Y bovino. Aparentemente, (ver Tabla 4) parecía haber un gran número de microsatélites disponibles para el estudio de la diversidad genética del cromosoma Y bovino. Sin embargo, la mayoría de los marcadores disponibles son monomórficos en machos, están situados en la zona PAR o presentan múltiples copias a la amplificación (Liu *et al.* 2002; Van Hooft *et al.* 2002) como luego se demostró en los artículos descritos en los puntos 3.2 y 3.3 del capítulo 3. Esto provoca que la lista de marcadores útiles para los propósitos de esta memoria se reduzca drásticamente. Asimismo, durante el desarrollo de este proyecto se han encontrado ciertas diferencias de comportamiento de algunos marcadores respecto a trabajos publicados con anterioridad. En consecuencia, el número de marcadores disponibles y útiles para propósitos filogenéticos sólo puede considerarse como muy reducido. En nuestras condiciones de trabajo expuestas en los artículos arriba mencionados, solo 5 microsatélites, que presentan en total 6 *loci*, resultaron útiles para establecer relaciones filogenéticas en el ganado bovino. Estos marcadores seleccionados coincidían básicamente con los utilizados por Ginja *et al.*, (2008). La información generada con estos microsatélites resultó, inevitablemente, limitada.

La búsqueda de SNP ante el desconocimiento de la secuencia del cromosoma Y es una tarea complicada. En el trabajo de Götherström *et al.*, (2005), se secuenciaron un total de 3,5 kb en intrones de genes del cromosoma Y bovino y sólo se identificaron cinco SNP, los únicos descritos hasta el momento en ganado bovino. Aunque no constituye la base de esta Memoria, durante el desarrollo del proyecto también se ha tratado de identificar variación genética asociada a secuencias codificantes localizadas en el cromosoma Y bovino. La estrategia más utilizada para la identificación de secuencias codificantes en el cromosoma Y bovino consiste en secuenciación comparativa, que no siempre produce los resultados esperados (Liu *et al.*, 2007b). Durante la estancia realizada en el laboratorio del Dr. Abel Ponce de León (*University of Minnesota*) se intentó la



caracterización de los genes XKRY, VCY y PRY (resultados no mostrados en esta memoria) utilizando cebadores desarrollados a partir de secuencias de la especie humana. A pesar de que esos genes están claramente localizados en la zona MSY del cromosoma Y humano (Skaletsky *et al.*, 2003) en ganado bovino se obtuvieron amplificaciones tanto en ADN de machos como en ADN de hembras. Los fragmentos amplificados presentaban la misma secuencia en ambos sexos y, a veces, la homología con la secuencia conocida para humanos era muy baja. Otra aproximación en este sentido se realizó en la estancia realizada en el laboratorio del Dr. Albano Beja-Pereira (*Universidade do Porto*), donde se intentó identificar SNP en 2 exones del gen UPS9Y a partir de la secuencia descrita para *B. taurus* (Van Laere *et al.*, 2008). Tampoco en este caso se identificó polimorfismo alguno (resultados no mostrados en esta memoria). Estos trabajos fallidos dan una idea de lo complejo que resulta la identificación de SNP informativos en el cromosoma Y bovino. En conjunto, la información acumulada permite considerar que: a) el polimorfismo de microsatélites es todavía una herramienta útil para el estudio de la diversidad del cromosoma Y bovino, aún en un momento en que los proyectos de secuenciación del genoma de las especies de mamíferos están alcanzando la madurez (Venter *et al.*, 2001; Human Genome Sequencing Consortium, 2001; Rhesus Macaque Genome Sequencing *et al.*, 2007; Bovine Hapmap Consortium, 2009; Tellman *et al.*, 2009); b) como ya había sugerido Hellborg *et al.*, en 2004, el cromosoma Y de mamíferos presenta baja diversidad nucleotídica, y c) es necesario identificar nuevas fuentes de información genética que permitan la caracterización de la diversidad en el cromosoma Y bovino.

Visto que el número de SNP disponibles en el cromosoma Y era muy escaso y que encontrar nuevos SNP es una ardua tarea, se planteó la posibilidad de testar la utilidad de los microsatélites que presentan múltiples copias a la amplificación (IMM) para realizar estudios poblacionales en ganado bovino. Hasta el momento, este tipo de microsatélites *multilocus* se han descrito en numerosas especies como

en *Homo sapiens*, *Felis catus*, equinos, *Bos taurus* y *Canis lupus* (Luo *et al.*, 2007; Wallner *et al.*, 2004; Liu *et al.*, 2003; Sundqvist *et al.*, 2001) pero nunca habían sido utilizados para ningún propósito práctico. La ventaja teórica de este tipo de marcadores consiste en que están íntimamente ligados a la naturaleza repetitiva del cromosoma Y, mecanismo que trata de paliar la degeneración del cromosoma Y debida a la falta de recombinación con el cromosoma X (Vallender y Lahn, 2004; Murphy *et al.*, 2006). Los microsatélites *multilocus* son el resultado de una PCR realizada con un solo par de cebadores con la cual se obtienen numerosas bandas. Los IMM presentaban distintos patrones de bandas entre animales de distintas poblaciones bovinas, es decir, eran polimórficos. Para poder utilizar estos marcadores se ensayó tratarlos como marcadores dominantes teniendo en cuenta la presencia o ausencia de bandas en los diferentes individuos. Cada banda sería un locus independiente de modo que asignando ceros y unos en base a la ausencia o presencia de bandas, cada individuo estaría caracterizado por un vector que constituiría la información a procesar. En todo caso, no todos los genotipos de los microsatélites *multilocus* o IMM son identificables con certeza lo que exigía requisitos adicionales para determinar su utilidad.

El trabajo realizado con este tipo de marcadores exigió profundizar, no sólo en su utilidad para el establecimiento de relaciones filogenéticas, sino en el conocimiento de su propia naturaleza. Para ello, se planteó estudiar de forma intensiva el único microsatélite utilizado hasta el momento en ganado bovino que presenta dos *loci*: el UMN0103. Este era el microsatélite de elección ya que en el concurrían diversas características: a) se puede considerar la hipótesis de que se trata de un IMM en estadios iniciales, con sólo dos bandas polimórficas en nuestros análisis, aunque estas 2 bandas pueden tratarse estadísticamente como dos *loci* clásicos (Mathias *et al.*, 1994); y b) en la bibliografía existían discrepancias sobre el número de bandas amplificables en este marcador: Ginja *et al.*, (2008) describen este marcador con un solo *locus* en *B. taurus* y dos en *B. indicus* mientras que en

trabajos anteriores de Liu *et al.*, (2002) y de Cai *et al.*, (2006) se describía como multicopia.

#### **4.2 Estructura de la información**

En la presente Memoria se ha entrelazado la información metodológica con la información de carácter filogenético e interés multidisciplinar.

En la sección 3.1 del capítulo 3 se propone una metodología que permita establecer con certeza la utilidad de los marcadores microsatélite para el estudio de la diversidad del cromosoma Y bovino. Esta metodología se demuestra en dos de los marcadores, (INRA124 e INRA126) supuestamente específicos del cromosoma Y bovino, que fueron ampliamente utilizados en estudios previos (Li *et al.*, 2007; Hanotte *et al.*, 2000; Edwards *et al.* 2000, 2007; Giovambattista *et al.* 2000). Esta información está contenida en el artículo: “Female segregation patterns of the putative Y-chromosome-specific microsatellite markers INRA124 and INRA126 do not support their use for cattle population studies”. De este trabajo metodológico se desprende la necesidad de una completa caracterización de los microsatélites asociados al cromosoma Y bovino así como la recomendación de tratar con cautela la información proporcionada con marcadores no suficientemente contrastados.

En el capítulo 3, sección 3.2 se determina que microsatélites son validos o no para testar la vía macho y se aplican estos a un amplio número de muestras de *B. taurus*. Esta información se recoge en el artículo: “Y-specific microsatellites reveal an African subfamily in taurine (*Bos taurus*) cattle”, en el que, como resultado más destacado se identificó un subgrupo Y2 en ganado bovino africano diferente a los cromosomas Y2 propios de los toros europeos.

En la sección 3.3 se demuestra, por primera vez, la capacidad de los marcadores IMM para establecer relaciones filogenéticas entre individuos y poblaciones bovinas, todo ello presentado en el artículo: “Multiple paternal origins of domestic cattle revealed by Y-specific interspersed multilocus microsatellites”. Los IMM permitieron diferenciar perfectamente los 3 haplogrupos identificados por Götherström *et al.*, (2005) mediante SNP, así como establecer tiempos de divergencia entre esos haplogrupos.

En el punto 3.4 se estudia en profundidad la estructura del microsatélite UMN0103. Este microsatélite es un ejemplo sencillo de microsatélite *multilocus*. Se recoge la información en la nota corta titulada “Genetic structure of the bovine Y- specific microsatellite UMN0103 reflects the genetic history of the species”, donde se vio que este marcador refleja la historia evolutiva de los distintos cromosomas Y descritos en ganado vacuno.

La información contenida en los artículos presentados en esta Memoria se discute en los siguientes puntos.

#### **4.3 Criterios de selección de marcadores específicos del cromosoma Y en ganado bovino.**

El primer paso fue una búsqueda bibliográfica de los marcadores microsatélites identificados en el cromosoma Y bovino y ovino (SRYM18). De forma rutinaria se testó su especificidad de macho. Algunos de esos marcadores, particularmente los BM861, INRA189, INRA124 e INRA126, ya habían sido utilizados previamente en estudios de diversidad (Li *et al.*, 2007) y también para diferenciar *B. taurus* de *B. indicus* (Hanotte *et al.*, 2000). Sin embargo, nuestros test de especificidad no siempre producían resultados compatibles con la información publicada previamente. Como ejemplo llamativo, los microsatélites INRA124 e INRA126 amplificaban en ADN de hembras bovinas con el mismo rango de tamaño que

presentaban los machos. Nuestros resultados eran compatibles con los publicados por Li *et al.*, (2007) en lo que respecta al INRA126 donde se sugería que este marcador podría estar situado en la zona PAR. Las discrepancias encontradas respecto de trabajos previos plantearon la necesidad de testar de forma exhaustiva los marcadores localizados en el cromosoma Y bovino previamente a su utilización en estudios de diversidad.

#### **4.3.1 Criterios generales**

Se definieron una serie de criterios que deben cumplir los microsatélites específicos del cromosoma Y (recogido en el artículo “Female segregation patterns of the putative Y-chromosome-specific microsatellite markers INRA124 and INRA126 do not support their use for cattle population studies”) que son los siguientes:

- Especificidad de macho
- Polimorfismo
- Segregación

##### **4.3.1.1 Especificidad de macho**

Este trabajo está enfocado a la historia de la línea masculina por lo que de todos los marcadores presentes en el cromosoma Y bovino, sólo son útiles a efectos filogenéticos aquellos situados en la zona MSY. Los marcadores presentes en la zona PAR, deben ser descartados. Con su utilización no se obtendría información exclusiva de la vía macho ya que recombinan con el cromosoma X. Para detectar la especificidad de macho los microsatélites deben testarse tanto en ADN de machos como en ADN de hembras. Las amplificaciones detectadas en ADN de hembras son frecuentes y pueden ser debidas a 4 razones (Liu *et al.*, 2003):

- Que se trate de un microsatélite situado en zona PAR.

- Que se produzcan amplificaciones inespecíficas, caracterizadas por el hecho de que el tamaño amplificado en ADN de hembras es distinto al amplificado en ADN de machos.
- Que en los cromosomas sexuales existan zonas similares debido a su origen común, caracterizado por que el tamaño de la amplificación es el mismo, aunque entre ellas no se produzca recombinación.
- Que existan múltiples copias de un marcador Y específico repartidas por el genoma, produciendo un mismo tamaño de la amplificación.

En ausencia de cartografía física, es difícil determinar la causa por la que un microsatélite localizado en el cromosoma Y amplifica en ADN de hembras. Como regla general, Liu *et al.*, (2003) sugieren que cuando se encuentra amplificación en ADN de hembras dentro del mismo rango de tamaños que amplifican en ADN de machos, esos microsatélites deben ser descartados.

Sea cual sea la causa de esa amplificación en ADN de hembras, a excepción de las amplificaciones inespecíficas, no debe descartarse la posibilidad de que ese microsatélite se encuentre situado en la zona PAR. Un 43% de los marcadores amplificaron en hembras, esto es, sólo 23 de los 44 marcadores testados amplificaron exclusivamente en ADN de machos o, como en el caso de los marcadores UMN2303 y UMN1514, amplificaron en ADN de hembras con productos de tamaños distintos de los obtenidos a partir de ADN de machos, con lo que resultarían marcadores útiles para indicar la vía masculina.

#### **4.3.1.2 Polimorfismo**

Un *locus* se considera polimórfico en una población, si podemos detectar más de un alelo. Generalmente el criterio más utilizado es el del 5%, es decir, un locus

será polimórfico si el alelo más común presenta una frecuencia menor al 95% en la población de estudio. En el estudio de Liu *et al.*, (2003) definían el polimorfismo en el cromosoma Y para la especie, incluyendo ADN de hembras lo que puede provocar la inclusión de polimorfismos de origen femenino. Si queremos estudiar exclusivamente las líneas paternas debe definirse el polimorfismo de un marcador dentro de la línea macho, sin incluir las líneas femeninas. Para detectar el polimorfismo de los marcadores MSY se testaron estos marcadores en un alto número de machos de distintas poblaciones de *B. taurus* y *B. indicus* con el fin de encontrar los diferentes alelos. Es de destacar, que de los 23 marcadores que cumplían los requisitos de especificidad de macho sólo 11 (47%) mostraron polimorfismo (en *B. taurus*).

Hay que destacar que se encontró un porcentaje elevado de marcadores *multilocus* y no todos han sido testados en este punto. De hecho los 44 marcadores específicos del cromosoma Y testados, 16 (36%) presentaron un comportamiento de este tipo.

#### **4.3.1.3 Segregación**

Este punto hace referencia a una de las características del cromosoma Y, concretamente de la zona MSY, que es la herencia de padres a hijos como un clon. Para la caracterización de las líneas paternas del ganado bovino sólo pueden ser útiles aquellos marcadores localizados en el cromosoma Y que se hereden de forma clónica de padres a hijos. Los microsatélites que cumplan con los requisitos de especificidad y polimorfismo anteriormente descritos deben testarse en familias de machos, donde se debe identificar el mismo genotipo entre un semental y sus descendientes masculinos.

#### **4.3.2 Criterios adicionales**

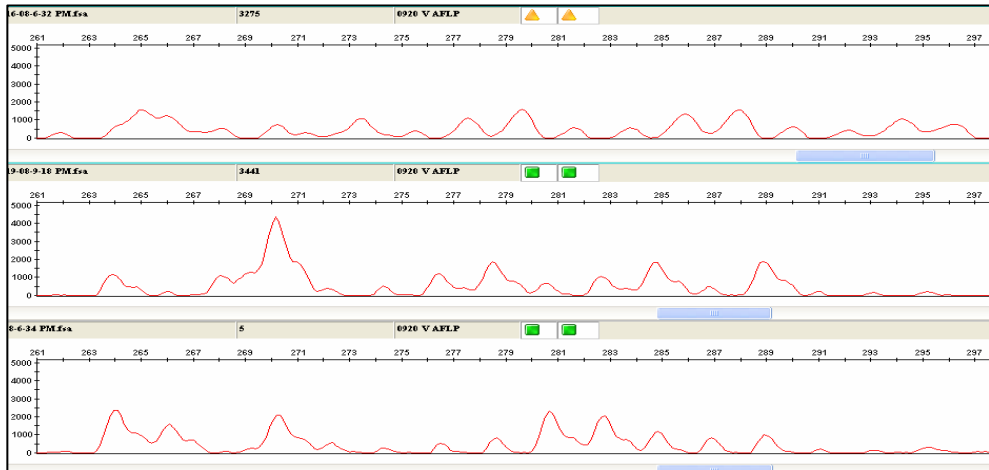
Los microsatélites localizados en el cromosoma Y, debido al estado de hemicigosis, se espera que presenten una sola banda. Sin embargo, es frecuente que presenten dos o más bandas. Como se ha determinado en la especie humana, este hecho puede darse incluso en poblaciones concretas aunque el comportamiento general sea el de amplificar una sola banda (Mathias *et al.*, 1994). Muchos de los microsatélites testados presentan un patrón de múltiples bandas ya reseñado en la bibliografía (Liu *et al.*, 2003). Estos marcadores son denominados marcadores microsatélite *multilocus* (*Interspersed Multilocus Microsatellites*; IMM). Debido a las características específicas de este tipo de marcadores, y su definición como marcadores dominantes, el testaje de su utilidad para estudios de diversidad debe incluir dos nuevos criterios adicionales a los anteriormente descritos: el Genotipado automatizable y la Repetibilidad (Información aportada en el artículo “Multiple paternal origins of domestic cattle revealed by Y-specific interspersed multilocus microsatellites”). Estas comprobaciones más rigurosas son comunes a otro tipo de marcadores definidos como dominantes, (como ejemplo AFLPs, Bonin *et al.*, 2007).

##### **4.3.2.1 Genotipado automatizable**

Este punto hace referencia a que el patrón de bandas que tiene que presentar un IMM debe ser un patrón de lectura sencilla y objetiva (Figura 12). Un IMM dado puede combinar diferentes bandas que presenten patrones de variación mono- y dinucleotídicos. Para genotipar microsatélites IMM cada banda se considera como un locus independiente y se asignan unos o ceros en base a la presencia o ausencia de bandas, respectivamente. El resultado obtenido para cada animal es un vector de ceros y unos similar, al resultado obtenido cuando se genotipan marcadores dominantes.



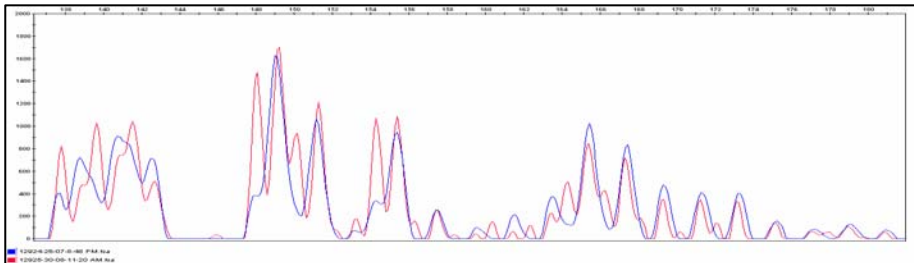
**Figura 12.** Marcador UMN0920, ejemplo de IMM de difícil genotipado. En el electroferograma se aprecia como no es posible definir con claridad cada banda.



#### 4.3.2.2 Repetibilidad

Esta característica hace referencia específicamente a los microsatélites IMM. La definición de estos marcadores como dominantes los hace más sensibles a los errores de genotipado (Pompanon *et al.*, 2007). Estos marcadores se ven más afectados que otros por las condiciones de PCR o migración con lo que presentan problemas de repetibilidad (Figura 13). Para testar este punto se deben genotipar los marcadores en los mismos individuos a partir de PCR distintas para comprobar la existencia o no de este problema.

**Figura 13.** Marcador UMN1201, ejemplo de IMM no repetible. En azul y rojo patrón de bandas del mismo individuo a partir de dos PCR realizadas en diferentes momentos.



En nuestro análisis, a cada muestra se le asignó un número con lo que población y especie no condicionasen el genotipado. Los IMM son particularmente sensibles a muestras de ADN de baja calidad por lo que muestras con estas características deben ser descartadas. Sistemáticamente, un 3% de controles ciegos se incluyeron en cada paso del procedimiento de genotipado. Cada muestra se genotipó 3 veces de forma independiente por la misma persona. Sólo se admitieron para posteriores análisis aquellos microsatélites que ofrecieron un patrón de bandas invariable.

#### **4.4 Aplicación de los criterios de selección de marcadores**

Siguiendo los criterios anteriormente mencionados un total de 44 marcadores de origen bovino y ovino fueron testados en *B. taurus* y *B. indicus*. Cuando un marcador no superó uno de los criterios no fue testado para el siguiente criterio. Se realizaron dos selecciones de marcadores.

##### **4.4.1 Primera selección de marcadores**

En la primera ronda de selección se descartaron los microsatélites que amplifican en hembras, los no polimórficos y los que presentan múltiples copias. Tras testar los microsatélites encontrados en la bibliografía para el cumplimiento los criterios fijados, se seleccionaron los marcadores BM861, INRA189, UMN0307, BYM-1 y UMN0103 que presentaba una o dos bandas al genotipado. Todos estos marcadores habían sido empleados con anterioridad en distintos estudios de diversidad o filogenia (Ginja *et al.*, 2008; Edwards *et al.*, 2000, 2007; Li *et al.*, 2007).

Otros marcadores utilizados ampliamente en otros estudios fueron descartados por no cumplir los criterios que se predeterminaron en el punto anterior. Ese es el caso del INRA124 donde se halló amplificación en hembras (Hanotte *et al.*, 2000; Edwards *et al.*, 2000, 2007; Giovambattista *et al.*, 2000) y también del INRA126

que, además de amplificar en hembras, no presenta una correcta segregación de padres a hijos. En el caso del marcador DYZ1 utilizados en otros estudios (Bradley *et al.*, 1994) no se logró amplificar por lo que no se obtuvieron resultados para este marcador y no fue utilizado en los posteriores análisis. El marcador UMN0504 (Ginja *et al.*, 2008) no fue utilizado por amplificar en hembras en el estudio presentado en esta Memoria, además de resultar monomórfico en el trabajo mencionado.

Los marcadores seleccionados se organizaron en dos PCR multiplex para facilitar el genotipado de los animales. Así en la PCR 1 se combinaron los microsatélites BM861, UMN0307 y BYM-1 y en la PCR 2 los microsatélites INRA189 y UMN0103. Estos marcadores presentan por lo general un bajo polimorfismo, presentando, de entre 2 a 9 alelos dependiendo del marcador. El microsatélite UMN0103 presentó dos *loci* con lo que se plantea un problema de genotipado. Se decidió seguir la Regla de Mathias (Mathias *et al.*, 1994; Balaesque *et al.*, 2006) que determina que la banda de menor tamaño es asignado como *locus* 1 y la banda de mayor tamaño es asignado como *locus* 2. Según Balaesque *et al.*, (2006), aunque pueden verse afectadas las frecuencias alélicas de cada población, no se ven afectadas las distancias genéticas entre poblaciones. Los resultados obtenidos coinciden con los de Balaesque *et al.*, (2006), lo que avala la posibilidad de utilizar la información proporcionada por el marcador UMN0103 como dos *loci* diferentes.

En ese sentido, es necesario señalar que este marcador ha presentado diferencias según el estudio realizado. Ginja *et al.*, (2008) sólo encontraba un *locus* por animal excepto para los animales *B. indicus*, que presentaban 2 *loci*. Cai *et al.*, (2006) describía este marcador como multicopia.

#### 4.4.1.1 Rendimiento filogenético de los microsatélites seleccionados

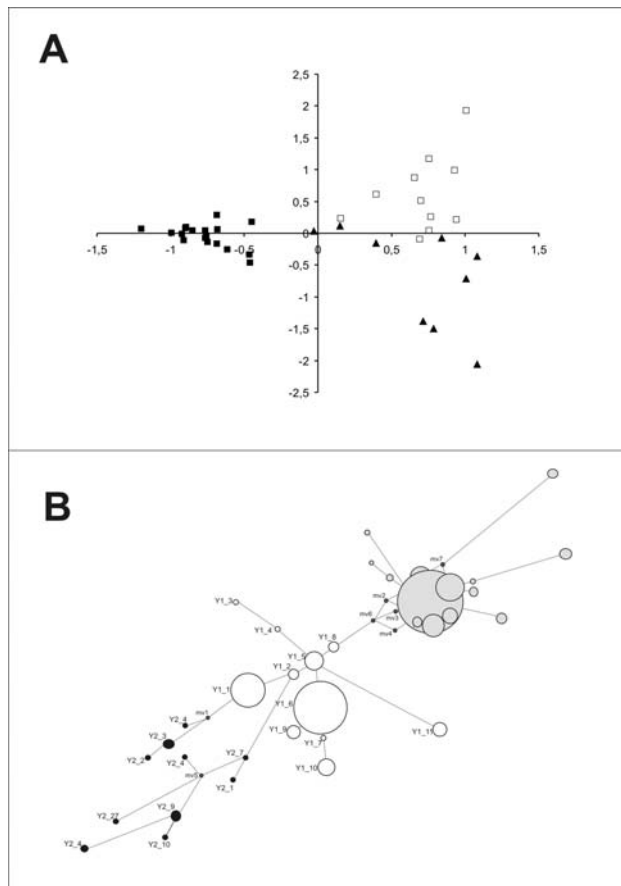
Los cinco marcadores (y seis *loci*) seleccionados se testaron en una amplio muestreo que incluía 608 animales pertenecientes a 45 poblaciones de *B. taurus*. De acuerdo con otros estudios que utilizaban este tipo de microsatélites (Li *et al.*, 2007; Ginja *et al.*, 2008; Kantanen *et al.*, 2009), el polimorfismo encontrado fue bajo, identificándose un total de 38 haplotipos de los que 14 resultaron únicos. El 62% de las muestras se concentraron en dos haplotipos mayoritarios.

El análisis estadístico de los resultados obtenidos indicó que los haplotipos identificados se organizaban en 2 grupos formados por: a) muestras procedentes de Europa Atlántica, Islas Británicas y poblaciones del África subsahariana; y b) muestras procedentes de razas de Europa continental y mediterránea y, mayoritariamente, de animales de África mediterránea. Esta situación contrastaba con la expuesta previamente por Götherström *et al.*, (2005) que sugerían que las razas de la Europa Atlántica e Islas Británicas (Y1) serían resultado de una introgresión de uro europeo en las poblaciones bovinas previamente domesticadas procedentes del Oriente Medio (Y2) y que serían mayoritarias en la Europa continental y mediterránea. El hecho de que el *B. taurus* del África Subsahariana se agrupara con individuos de la Europa Atlántica apoyaría las sugerencias de Svenson y Götherström (2008) y Bollongino *et al.*, (2008), que utilizando muestras de ADN antiguo rebatieron la tesis inicial de Götherström *et al.*, (2005), sugiriendo que la situación presente era resultado de un proceso de deriva.

Al contrastar los haplotipos identificados con los SNP zfy9 y zfy10 de Götherström *et al.*, (2005), los incluidos en el grupo b) anterior se asignaron al haplogrupo Y2, mientras que los incluidos en el grupo a) se asignaron al haplogrupo Y1 excepto los procedentes de toros del África Subsahariana y un animal de Marruecos, que se asignaron al haplogrupo Y2. En la Figura 14 puede comprobarse como los análisis de correspondencia y NETWORK llevados a cabo con el análisis

conjunto del polimorfismo de microsatélites y SNP, no variaba de forma sustancial la información proporcionada por los microsatélites, confirmando la situación inicial.

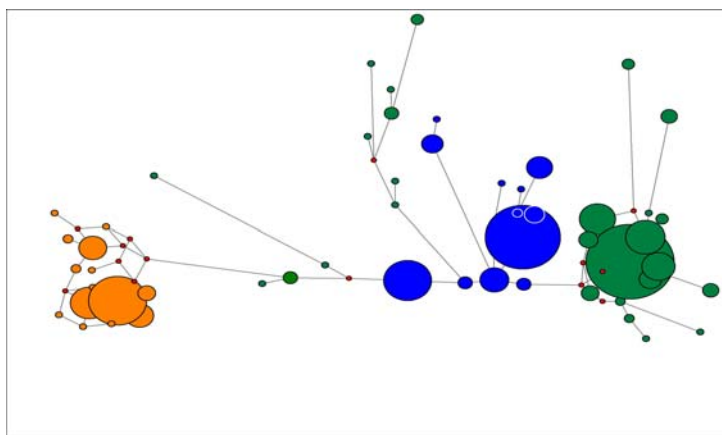
**Figura 14.** Representación grafica de la variabilidad genética entre los haplotipos Y1 e Y2. En A) Dispersión de los 38 haplotipos calculados con Análisis de correspondencia, haplotipos Y1 representados con triángulos negros, Y2 con cuadrados negros a excepción de los Y2 africanos representados por cuadrados blancos. En b) Red network construida con el programa NETWORK 4.5.2. Los haplotipos Y1 se representan con círculos blancos, Y2 con círculos grises a excepción de los Y2 africanos representados por círculos negros. En la figura se puede apreciar como los haplotipos Y2 africanos clusterizan con los haplotipos Y1. El tamaño de los círculos es proporcional a la frecuencia del haplotipo.



En conjunto, este análisis ofreció dos conclusiones de interés:

Como cuestión de interés general, la batería de marcadores ensayada permitió, por primera vez, la diferenciación de un subgrupo africano dentro del haplogrupo Y2. Asumiendo que los uros africanos (*B. primigenius opisthonomus*) deberían ser genéticamente similares a los domesticados en Oriente Medio (Bradley y Magee, 2006), nuestros resultados apoyarían una posible domesticación de *B. taurus* en África. La posibilidad de esta domesticación fue sugerida primero a partir de evidencias arqueológicas (Meadow, 1993; Wendow y Schild, 1994). Análisis de microsatélites en muestras europeas y africanas sugieren esta misma hipótesis (MacHugh *et al.*, 1997). La presencia de ADN mitocondrial diferente (el grupo T1) y la presencia de haplotipos únicos en África (Troy *et al.*, 2001; Loftus *et al.*, 1994a; Bradley *et al.*, 1996) sugieren que el ganado vacuno africano no es un simple muestra de aquel domesticado en el Creciente Fértil (Bradley y Magge, 2006). La presencia de haplotipos singulares en el cromosoma Y de *B. taurus* africanos, ahora identificada, supone un importante apoyo a la hipótesis de una posible domesticación africana.

**Figura 15.** Análisis NETWORK en el que se incluyeron muestras de *B. taurus* y *B. indicus*. Se generaron 55 haplotipos con el set de 5 microsatélites al añadir un total de 138 muestras originarias de África y Asia de *B. indicus* a las muestras analizadas en el artículo Pérez-Pardal *et al.* 2010a. En azul Y1, verde Y2 y naranja Y3 (Datos no mostrados en esta Memoria).



Aunque en el artículo no se utilizaron animales cebú, análisis complementarios se realizaron con poblaciones *B. indicus*. Así estos individuos fueron genotipados con los cinco microsatélites e incluidos a la red Network en color naranja (Figura 15), todos los individuos cebú fueron contrastados con los SNP de Götherström. Sin embargo, pese a que discrimine bien entre *B. indicus* y *B. taurus* desde el punto de vista metodológico, los análisis realizados permiten evidenciar las limitaciones de esta batería de microsatélites: a) cuando se utilizan sin la aplicación de los SNP no es posible diferenciar correctamente los haplogrupos Y1 e Y2, que parecen tener importancia filogenética; b) por si mismo, la batería de microsatélites utilizada no permite diferenciar el Y1 del Y2 africano; y c) la información proporcionada no permite calcular tiempos de divergencia entre haplogrupos. La batería de microsatélites ensayada se compone de marcadores de evolución rápida (Goldstein *et al.*, 1995). En esta situación, con un limitado número de microsatélites disponibles es probable la presencia de homoplasias que impidan una correcta diferenciación de los haplogrupos. El trabajo aquí presentado pone en evidencia las limitaciones de usar sólo los SNP identificados por Götherström *et al.*, (2005), que permiten definir los haplogrupos Y1 e Y2 pero no permite detectar la familia africana del haplogrupo Y2.

Las carencias detectadas en esta colección de marcadores utilizados, hacen necesario el desarrollo de nuevas herramientas para el análisis del cromosoma Y bovino.

#### **4.4.2 Segunda selección de marcadores. Desarrollo de nuevas herramientas**

En la primera selección de marcadores se habían descartado los microsatélites que amplificaban en hembras por su posible situación en la zona PAR, los microsatélites monomórficos por su carencia de información y los *multilocus* (IMM) que nunca antes se habían utilizado como tal. Los marcadores IMM, como ya se explicó anteriormente, son el resultado de una PCR amplificada con un solo

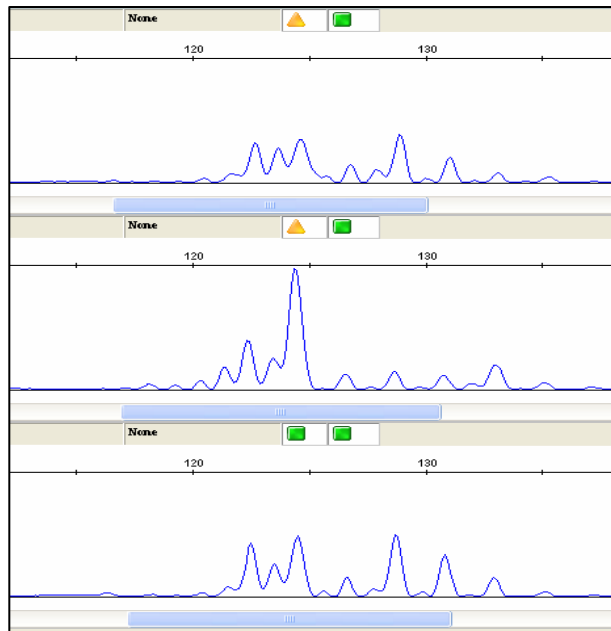
par de cebadores y con la cual se obtienen numerosas bandas. Aunque estos marcadores habían sido descritos en numerosas especies de mamíferos nunca habían sido empleados con propósitos filogenéticos. Sin embargo, existe una importante razón teórica que incita a testar la utilidad de este tipo de marcadores: su definición está íntimamente ligada a la naturaleza repetitiva del cromosoma Y (Vallender y Lahn, 2004; Murphy *et al.*, 2006) y, en consecuencia, pueden ofrecer información singular que permita caracterizar las fuerzas evolutivas que influyen en esta parte del genoma. A pesar de ello, existe un problema en su definición y, por lo tanto, en la forma de proceder a la codificación de la información que proporcionan, así como al establecimiento de criterios adicionales para la selección de IMM y para su genotipado.

Un aspecto fundamental para poder analizar estos marcadores consiste en la forma de codificar su información. Se decidió tratar cada banda como un *locus* independiente y con carácter dominante. Para ello se tiene en cuenta la presencia o ausencia de cada banda asignando un uno a la presencia de la banda y un cero a su ausencia. Cada individuo está, por lo tanto, caracterizado por un vector de ceros y unos que va a ser la información que se va a procesar. Este tipo de codificación tiene la ventaja de la experiencia acumulada en la comunidad científica en el tratamiento de marcadores dominantes (ver como ejemplo Bonin *et al.*, 2007). En todo caso, es necesario señalar que con este método no se está teniendo en consideración toda la capacidad informativa de los IMM. Teóricamente, es posible que dos muestras que presenten el mismo haplotipo para un IMM dado varíen en número de copias de un *locus* dado, un ejemplo de esto es el marcador INRA057 (Figura 16). En hombre y ratón, la variación en el número de copias en segmentos de ADN somático se considera una importante fuente de variación genética que afecta a la expresión genética y a la variabilidad fenotípica (Freeman *et al.*, 2006; Perry *et al.*, 2006; Egan *et al.*, 2007; Stranger *et al.*, 2007). Sin embargo, la



identificación y utilización de este tipo de variabilidad superaba los objetivos del presente proyecto.

**Figura 16.** *Microsatélite INRA057, ejemplo de marcador que presenta el mismo haplotipo de ceros y unos presentando polimorfismo de copia. A simple vista se aprecia que el electroferograma del medio presenta diferencias respecto de los otros dos, pero la codificación sería la misma para los tres individuos.*



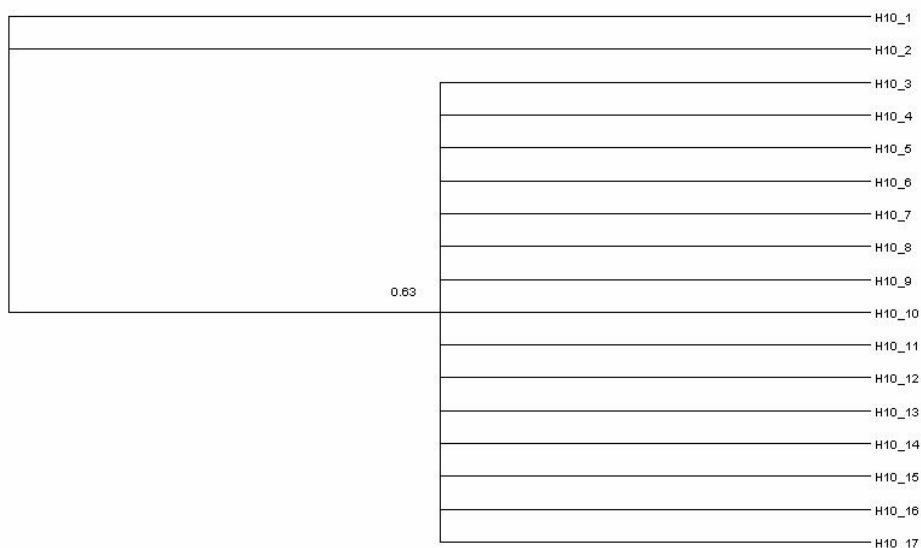
Ante la complejidad de genotipado que presentan estos marcadores, la selección de marcadores informativos se centró en cinco de ellos (UMN0920, UMN1201, UMN2713, UMN2303 y UMN2405) que fueron escogidos en base al polimorfismo que presenta el patrón de bandas. En conjunto, los cinco IMM testados presentaron 136 loci de los que 83 (61%) eran polimórficos. Se descartaron los marcadores UMN0920 y UMN1201 por no cumplir los criterios de genotipado automatizable y repetibilidad respectivamente. El marcador UMN0920 presenta un patrón de bandas no claramente definido, por lo que la automatización del genotipado se hace imposible (Figura 10). En cuanto al marcador UMN1201 se

observó que su patrón de bandas no era repetible ya que con las mismas condiciones de PCR el patrón de bandas presentaba variaciones que imposibilitan un correcto genotipado (Figura 11). Los marcadores UMN2713, UMN2303 y UMN2405 cumplieron todos los requisitos requeridos.

#### 4.4.2.1 Rendimiento filogenético de los IMM

Se genotiparon un total de 453 animales pertenecientes a 52 poblaciones tanto de *B. indicus* como de *B. taurus* con los marcadores UMN2713, UMN2303 y UMN2405. Se realizaron análisis estadísticos de cada marcador por separado y de todas las combinaciones posibles. Esto permitió comprobar que el marcador UMN2713 no resulta informativo careciendo de capacidad discriminante (Figura 17), por lo que se descartó para posteriores análisis (datos de genotipado no mostrados en esta memoria).

**Figura 17.** Árbol MrBayes realizado con los 17 haplotipos resultantes de genotipar el marcador UMN2713 en una batería de muestras de *B. taurus* y *B. indicus*.



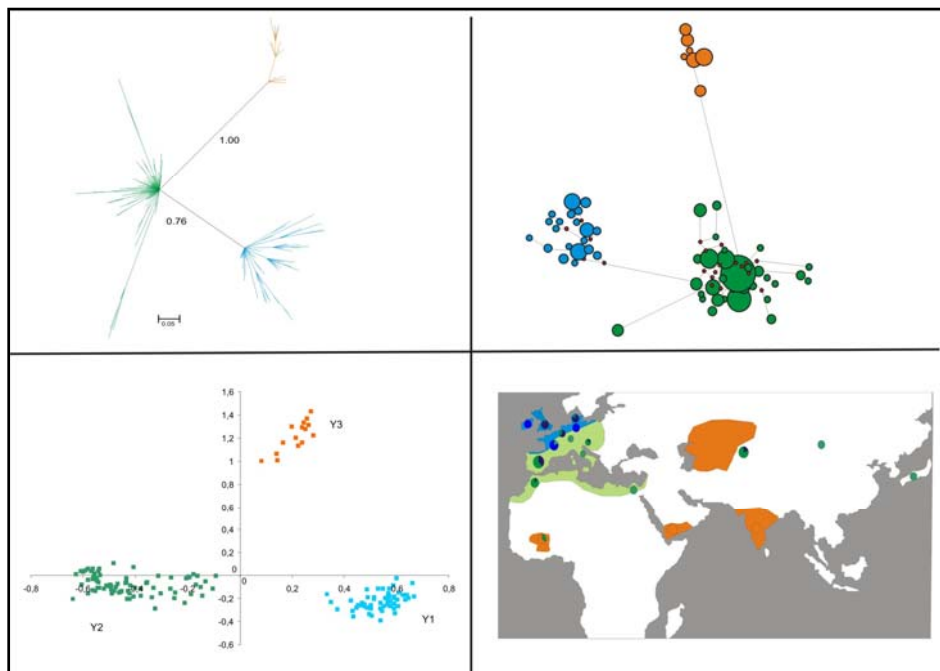
Los marcadores seleccionados (UMN2303 y UMN2405) presentaban 66 *loci* distintos de los que 48 (72%) eran polimórficos. En total, los dos marcadores utilizados recogían el 58% del total del polimorfismo de los cinco IMM testados. La información de estos dos marcadores permitió identificar 146 haplotipos de los que 87 resultaron únicos. Se analizaron los distintos haplotipos con el programa MrBayes, NETWORK y se hizo un análisis de correspondencia (Figura 18) identificándose tres grupos. De modo análogo a lo realizado con los haplotipos obtenidos mediante marcadores de comportamiento clásico, los tres grupos identificados mediante los IMM se contrastaron con los SNP de Götherström *et al.*, (2005). Los tres grupos coincidieron con los haplogrupos Y1, Y2 e Y3. La información proporcionada por los IMM seleccionados permitía, además: a) caracterizar el haplogrupo Y1 como el de mayor diversidad haplotípica lo que no permite descartar que sea el resultado de una introgresión de uro europeo en el ganado doméstico; b) confirmar la existencia de una subfamilia haplotípica dentro del haplogrupo Y2; esta subfamilia, compuesta por animales africanos, presentaba la mayor diversidad haplotípica relativa apoyando la hipótesis de la posible domesticación de *B. taurus* en África; c) identificar una mayor estructura genética dentro del haplogrupo Y2 sugiriendo posibles procesos locales de introgresión de uro sobre ganado doméstico; y d) caracterizar el haplogrupo Y3 como el menos diverso. Este hecho puede estar condicionado por el pequeño número de muestras *B. indicus* analizadas, sería compatible con un solo proceso de domesticación.

Desde el punto de vista metodológico la utilización de los IMM seleccionados presenta las siguientes ventajas:

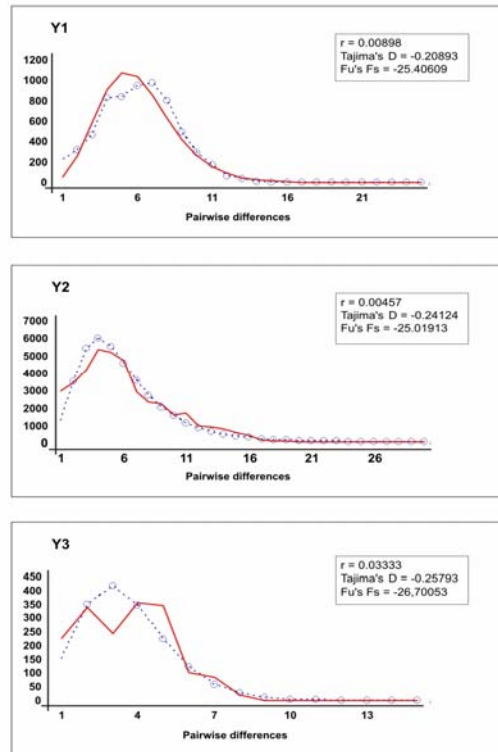
- son capaces de identificar la mayor fuente de variabilidad genética neutra conocida hasta el momento en el cromosoma Y bovino.

- Esta variabilidad genética está íntimamente ligada a la naturaleza del cromosoma Y.
- Los resultados obtenidos son concordantes con los obtenidos mediante SNP, permitiendo diferenciar los haplogrupos Y1, Y2 e Y3, y confirmando que los SNP identificados por Götherström *et al.*, (2005) tienen sentido filogenético.
- La información proporcionada por los IMM permite, asimismo, identificar estructura genética dentro de haplogrupos.

**Figura 18.** En el primer recuadro árbol filogenético construido con el programa MrBayes a partir de los 146 haplotipos encontrados. En el segundo recuadro análisis el network y en el tercero un análisis de correspondencias. En el mapa se representa la frecuencia de haplogrupos por país. En azul se identifica el Y1, en verde el Y2 y en naranja el Y3.



**Figura 19.** Distribución Mismatch y raggedness ( $r$ ),  $D$  de Tajima  $D$  y  $F_s$  de  $F_u$ , test de neutralidad estadísticos para cada haplogrupo definido (Y1, Y2 and Y3). La línea roja y la línea de puntos representan la distribución Mismatch esperada y la observada respectivamente.



Análisis complementarios, asumiendo un modelo de mutación IAM (*Infinite Allele Model*), incluyeron una “mismatch distribution” así como el cálculo de los estadísticos, raggedness ( $r$ ),  $D$  de Tajima y  $F_s$  de  $F_u$  entre los tres haplogrupos calculados con el programa Arlequin 3.11 (Excoffier *et al.*, 2005). La distribución *mismatch* cuenta el número de diferencias entre cada par de secuencias en una población (Rogers, 2002) y permite comparar la distribución esperada con la distribución observada. Es usualmente multimodal en muestras tomadas de una población en equilibrio demográfico, pero es generalmente unimodal en poblaciones que han sufrido una reciente expansión demográfica (Schneider *et al.*,

2000). Como se puede observar en los histogramas de la Figura 19 este análisis realizado en los tres haplogrupos, Y1, Y2 e Y3, resultó consistente con tres eventos de expansión de la población. El valor de *raggedness* fue no significativo, indicando que la distribución observada no se desvía de un modelo de expansión poblacional repentino. El estadístico D de Tajima fue negativo para los tres haplogrupos, como era de esperar cuando la estructura genética esta influida por una rápida expansión, y no significativo demostrando que no hay desviación de las expectativas de neutralidad. El estadístico  $F_s$  de Fu, que se corresponde con la distribución haplotípica, fue significativo para los tres haplogrupos indicando desviación del equilibrio mutación-deriva.

Todos los análisis efectuados permiten considerar a los IMM como una herramienta válida para el estudio de diversidad genética en el cromosoma Y bovino. De forma adicional, estos marcadores permitieron, por primera vez, estimar tiempos de divergencia entre haplogrupos del cromosoma Y bovino. La forma más correcta de tratar los IMM, codificados como marcadores dominantes, es considerarlos marcadores con una baja tasa de mutación. Sin embargo, aún cuando los IMM se consideraran marcadores de evolución intermedia o alta (Tabla 4), se obtuvieron tiempos de divergencia entre *B. taurus* y *B. indicus* (entre 75.292 y 658.780 años) siempre compatibles con una separación genética previa a la domesticación. En 1997, MacHugh *et al.*, usando microsatélites autosomales, habían estimado la divergencia entre *B. taurus* y *B. indicus* de 610.000 a 850.000 años usando *Bison sp.* como referencia. Utilizando ADN mitocondrial se estableció que el tiempo de divergencia entre cebú y toro variaría entre 200.000 y 1.000.000 de años (Loftus *et al.* 1994a). Otra estimación utilizando ADN mitocondrial estimó la divergencia entre estas dos especies entre 1.700.000 a 2.000.000 años, en este caso la discordancia entre líneas maternas y paternas podría estar causada por factores como un ratio de migración de machos-hembras descompensado y la poligamia (Cavalli-Sforza y Minch, 1997; Oota *et al.*, 2001; Seielstad *et al.*, 1998).

**Tabla 4.** Tiempos de divergencia y desviación estándar (d.e.) calculados entre haplotipos ancestrales (Pérez-Pardal *et al.*, 2010) usando el estimador  $\rho$  calculado mediante el programa NETWORK 4.2 (Forster *et al.*, 1996) Se han utilizado tres tasas de mutación diferentes: a) rápida: 0,0009/generación (Zhivotovsky *et al.*, 2004); b) intermedia: 0,0008/generación (Zhivotovsky *et al.*, 2004); y c) baja: 0,00018/generación (Forster *et al.*, 2000)

Haplotipo ancestral	J12-J19	J12-J24	J19-J24
Haplogrupo	Y3-Y2	Y3-Y1	Y2-Y1
$\rho$ (d.e.)	14,00 (0,71)	24,50 (1,00)	14,50 (0,71)
Años (d.e.)			
Tasa de mutación rápida	75292,0 (3802,8)	131761,0 (5378,0)	77981,0 (3802,8)
Tasa de mutación intermedia	84700,0 (4278,0)	148225,0 (6050,0)	87725,0 (4278,0)
Tasa de mutación lenta	376446,0 (19013,4)	658780,5 (26889,0)	389890,5 (19013,4)

#### 4.5 Caracterización del microsatélite UMN0103

En el planteamiento inicial de este trabajo no estaba incluido el estudio a una mayor profundidad del microsatélite UMN0103. Esta nueva exigencia radica en dos cuestiones fundamentales, la presencia de dos bandas en este marcador podría indicar que se trate de un microsatélite IMM en estadios iniciales con lo que su estudio podría llevar a comprender un poco mejor este tipo de microsatélites. La otra necesidad se debe, al igual que sucedía con los microsatélites INRA124 e INRA126, a diferencias detectadas entre los resultados aportados en esta Memoria y la bibliografía disponible para este marcador. Este marcador se utilizó en el capítulo 3 sección 3.2 para caracterizar el ganado bovino por la vía paterna aplicando la regla de Mathias para genotiparlo (Mathias *et al.*, 1994), por lo que se creyó conveniente dilucidar cualquier duda relacionada con este microsatélite.

El análisis de secuenciación reveló que en *B. taurus* los dos *loci* presentaban la misma secuencia, mientras que en *B. indicus* la presencia de una mutación en la zona de repetición del microsatélite podía discriminar entre los dos *loci*. En

ninguna de las dos especies fue posible el diseño de cebadores para caracterizar cada loci por separado.

Respecto a las diferencias halladas con las referencias bibliográficas estas pueden ser debidas a dos causas:

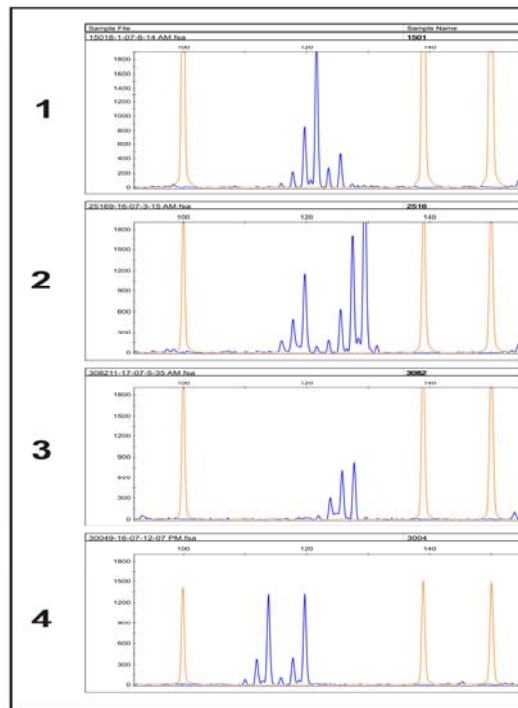
- a) En el caso de los estudios de Cai *et al.*, (2006) y Liu *et al.*, (2003) las diferencias pueden ser debidas al diferente método de genotipado, en estos estudios se usaban geles de poliacrilamida, método mucho más impreciso que el uso de secuenciadores automáticos, con lo que la presencia de los dos loci y las bandas fantasmas podría interpretarse fácilmente como multicopia.
- b) Las diferencias con el trabajo de Ginja *et al.*, (2008) estarían encaminadas a diferencias en la zona de hibridación de los cebadores en uno de los dos loci. Ginja *et al.*, (2008) utilizan PCR más estrictas. En esas condiciones uno de los loci se pierde en *B. taurus* pero no en *B. indicus*. En nuestras condiciones de PCR obteníamos los dos loci en las dos especies, con una salvedad, en *B. taurus* uno de los loci presentaba una mayor altura de pico que el otro, siendo del mismo tamaño en *B. indicus*.

Cuando se relacionó la morfología del microsatélite con la pertenencia a un determinado haplogrupo del cromosoma Y se observó con una PCR estándar que el locus con menor número de repeticiones se hallaba presente pero en una intensidad mayor que el otro locus, en condiciones estrictas los animales Y1 presentaban solo el locus de menor número de repeticiones. Lo mismo sucedía al locus de mayor número de repeticiones para los animales Y2. Dentro de los animales Y2 se encontraron animales que solo presentaban un locus. Estos animales en los análisis realizados con los IMM y el set de 5 microsatélites constituían una subfamilia dentro del haplogrupo Y2. En estos animales no fue posible determinar si hay al menos dos loci del mismo número de repeticiones o por el contrario había perdido



un *locus*. Basándonos en estas cuestiones podemos asegurar que este marcador resume bien la historia evolutiva de los machos bovinos. Como se aprecia en la Figura 20 se pueden distinguir los haplogrupos Y1, Y2 e Y3 además de la subfamilia Y2 africana.

**Figura 20.** Electroferogramas de un animal Y1 en 1, Y2 en 2, Y2 africano en 3 y un Y3 en 4. Se puede apreciar como el locus 2 en animales Y1 (en 1) y el locus 1 en los Y2 (en 2) presentan una menor altura de pico, mientras que los animales Y3 (en 4) presentan la misma altura de pico para los dos loci. En 3 se puede ver como un animal africano presenta un único locus.



La naturaleza repetitiva del cromosoma Y se postula como un mecanismo para mantener los genes específicos del cromosoma Y en un ambiente no recombinante que podría erosionar al cromosoma Y con un número elevado de mecanismos entre los que se incluye el estado de hemizigosis (Vallender y Lahn, 2004; Murphy *et al.*, 2006). Hay dos mecanismos principales de multiplicación del ADN: la

recombinación y la transposición de elementos móviles. Mediante eventos de recombinación internos, como sobrecruzamientos desiguales y conversión génica se generan homólogos en tarden (Van't Hof *et al.*, 2007). En cuanto a la transposición de elementos móviles, una gran cantidad de microsatélites en el genoma humano se han originado a partir de colas polyA de retrotransposones, principalmente elementos *Alu* y LINE-1. En un estudio se encuentra que un 72% de los microsatélites están relacionados con el final 3' de retrotransposones (Nadir *et al.*, 1996). Otro estudio sugiere que el 54% de los microsatélites del cromosoma Y se originaron por mediación de retrotransposones (Kayser *et al.*, 2004). Respecto a estos dos mecanismos de generación de nuevo material genético no hay consenso si se trata de procesos redundantes y separados o si por el contrario se trata de un mecanismo integrado (Van't Hof *et al.*, 2007).

Comparamos la secuencia obtenida para el UMN0103 con secuencias depositadas en el GeneBank y encontramos muchos clones que mostraban una identidad con nuestra secuencia del 99%. Analizamos algunos de esos clones en búsqueda de elementos móviles y se vio que la secuencia homóloga estaba rodeada de elementos móviles, como SINE (*Short interspersed repetitive elements*), LINE (*Long interspersed repetitive elements*) y LTR (*Long terminal repeats*), lo que podría explicar la duplicación del *locus* en este marcador.

#### **4.6. Repercusión, implicaciones y perspectivas futuras de este trabajo**

Este trabajo aquí presentado debería influir en un amplio rango de campos: Uno de ellos alude a la historia del ganado bovino. Se ha descrito por primera vez la estructura genética de los haplogrupos Y1, Y2 e Y3 confirmando la existencia de una subfamilia africana en el Y2. Esto podría revelar una posible domesticación de *B. taurus* en África. Y además se ha estimado la divergencia entre los distintos haplogrupos por vía paterna, encontrando concordancia con los tiempos descritos para la línea materna. La historia genética dibujada por los análisis del ADN

mitocondrial presenta concordancia con la información obtenida con los IMM. Esta cuestión resulta sorprendente ya que en otras especies, como oveja o cabra, la información obtenida por la vía masculina muestra un escenario diferente a la obtenida por la vía femenina (Pidancier *et al.*, 2006). La falta de coincidencia entre información genética masculina y femenina se explica asumiendo que la domesticación y la posterior dispersión fue diferente para cada uno de los sexos. El ganado vacuno presenta una baja capacidad de dispersión, por lo que su dispersión, después de la domesticación, fue muy lenta y probablemente dependiente del manejo del hombre.

Para conocer mejor la historia del ganado bovino deberían analizarse con esta colección de marcadores animales procedentes del Creciente Fértil, el centro de domesticación descrito para *B. taurus*. Una tarea pendiente de esta memoria es testar los cinco microsátélites y los dos IMM en individuos procedentes del Creciente Fértil, para dilucidar si realmente existió domesticación en África. Sería interesante e importante testar estos marcadores (tanto los IMM como la colección de 5 microsátélites) en muestras de ADN antiguo. La longitud de los fragmentos amplificables depende de las condiciones de preservación del resto antiguo (Burger *et al.*, 1999). Son muchos los autores que señalan que el ADN antiguo suele encontrarse degradado en fragmentos, de uno a unos pocos cientos de bases (Pääbo, 1985a, b; Hagelberg y Sykes, 1989; Hänni *et al.*, 1990; Rogan y Salvo, 1990; Goloubinoff *et al.*, 1993; Handt *et al.*, 1994) por lo que no es posible la amplificación por *PCR* de fragmentos de ADN superiores a 150 pares de bases (pb) en tejidos blandos (Pääbo, 1989). Este límite es sensiblemente superior en tejidos sólidos, como hueso o diente, siendo de 300 pb en ADN mitocondrial y de poco más de 100pb en material nuclear (Hagelberg *et al.*, 1994a; Hauswirth *et al.*, 1994). Testar este tipo de marcadores *multilocus* es una tarea que se presenta muy complicada, si no imposible, debido a la fragilidad del ADN antiguo (Pääbo *et al.*, 2004; Willerslev y Cooper, 2005).

Quedan por testar más marcadores IMM en Bovino, tarea que no se ha realizado para esta memoria debido a la cantidad de tiempo y complejidad que conllevan este tipo de marcadores. Sería interesante comprobar el conocimiento que aportan a la información ya revelada con los microsatélites utilizados en la presente memoria. Otra tarea pendiente es secuenciar los microsatélites monomórficos en un amplio rango de animales pertenecientes a diferentes haplogrupos para tratar de encontrar SNP. La misma tarea se debería hacer con los microsatélites polimórficos ya que muchos de ellos presentan, en animales poco relacionados, el mismo tamaño alélico. Esta coincidencia de tamaños podría deberse a homoplasias por lo que se podrían encontrar SNP en esas secuencias que permitiesen diferenciar esos alelos que presentan el mismo tamaño pero una diferente historia evolutiva.

Otro campo que se debería ver beneficiado por el conocimiento adquirido con esta memoria es el de la conservación, ya que por primera vez se ha estudiado la diversidad existente en los tres haplogrupos definidos por vía paterna. El conocimiento la variabilidad existente es el paso previo para proteger la variabilidad.

Esta Memoria pretende contribuir a que se conozca un poco más el cromosoma Y bovino, así como debería influir en la aplicación de los marcadores tipo IMM en otras especies. Con este trabajo se demuestra el valor informativo de algunos marcadores IMM, abriendo una puerta al uso de estos marcadores que se han descrito en numerosas especies incluyendo el hombre (*Homo sapiens*), caballo (*Equus caballus*), burro (*Equus asinus*), gato (*Felis silvestris catus*), cabra (*Capra hircus*) y oveja (*Ovis aries*) (Butler *et al.*, 2005; Luo *et al.*, 2007; Wallner *et al.*, 2004; Liu *et al.*, 2003; Sundqvist *et al.*, 2001; Pérez-Pardal *et al.*, no publicado). Debido a la carencia de marcadores existentes para analizar el cromosoma Y, los

IMM se revelan en este estudio como una fuente de información que no debería ser despreciada. Dependiendo del marcador, el genotipado tendrá un mayor o menor grado de complicación. Los marcadores descritos en esta memoria son bastante complejos debido al alto número de bandas que presentan y a que no hay ningún programa diseñado especialmente para genotiparlos. En otras especies se han encontrado marcadores IMM que presentan un menor número de bandas lo que los hace más fácilmente genotipables, como es el caso del marcador DSY464 en humanos (Butler *et al.*, 2005).

El uso de los marcadores IMM, pese a todas sus aportaciones, dista de ser idílico. Una de las principales desventajas que presentan estos marcadores es el esfuerzo de genotipado que, aunque varía dependiendo del número de locus que presente, es mucho mayor que para un microsatélite tradicional. Se han analizado IMM que constaban desde 4 loci (en *O. aries*, resultados propios no incluidos en esta memoria) hasta los descritos en esta memoria que presentan un total de 33 loci. Además no todos los IMM resultan informativos o utilizables debido a su codificación imposible.

El modelo de mutación de los IMM no se conoce. Durante los últimos años se han publicado numerosos estudios encaminados a la descripción de aproximaciones teóricas que puedan explicar el modelo evolutivo de los microsatélites. Uno de los modelos posibles es el conocido como *SMM* (*Stepwise Mutation Model*, o modelo mutacional paso a paso) (Kimura y Ohta, 1978). Este modelo se concibió inicialmente para explicar la regularidad en la distribución de las frecuencias alélicas, y asume que un alelo sólo puede mutar a sus dos estados adyacentes por pérdida o adición de una repetición. Un modelo alternativo, el IAM (*Infinite Allele Model* o modelo de alelos infinitos) (Kimura y Crow, 1964), propone que todas las mutaciones son posibles, sin que existan restricciones en el número de unidades repetidas mutadas, cada mutación da lugar al surgimiento de

un nuevo alelo que no existía en la población. En el modelo SMM, al contrario de lo que sucede en el IAM, los alelos pueden mutar hacia estados alélicos ya presentes en la población y de ahí generar lo que se denomina homoplasia. Para un marcador genético, la homoplasia ocurre cuando diferentes copias de un locus son idénticas en estado pero no idénticas por descendencia. En el IAM las diferencias de tamaño no son informativas respecto al grado de diferenciación entre poblaciones, mientras que el SMM las diferencias de tamaño entre alelos son una medida de diferenciación entre las poblaciones. En este sentido asumiendo el modelo SMM se pueden hacer cálculos dentro de poblaciones y asumiendo el IAM permitiría hacer cálculos entre especies. En lo que respecta al modelo de mutación de los IMM, se deben de desarrollar más análisis para ajustar mejor el modelo mutacional.

Otro punto que hay que esclarecer es la tasa de mutación que hay que aplicar a de este tipo de marcadores. En nuestros análisis asumimos que estos microsatélites son de baja evolución, aunque también analizamos los datos considerando que podrían tener una tasa de evolución media y alta (Tabla 4). De todos modos nunca se han hecho estudios para determinar la tasa de mutación de estos marcadores por lo que sería necesario que se analizasen un número elevado de unidades familiares para estimar su tasa de mutación.

También es necesario el refinamiento del método de genotipado propuesto en esta memoria en el que no se tiene en cuenta el número de copias por *loci*. Por el momento no hay una forma objetiva de cuantificar el número de copias, produciéndose una infraestima de la diversidad. Un error que se comete en sentido contrario es la posible consideración de las bandas fantasmas o *stater bands* como *loci*, una posible forma de evitar este problema va encaminada al trabajo que realizado con el artículo en el que se caracteriza el microsatelite UMN0103, donde este marcador era secuenciado para intentar separar los dos *loci*. En este trabajo no

fue posible separar mediante PCR los dos *loci*, pero quizás en otros IMM si sea posible lograr este objetivo. Mediante la separación de los distintos *loci* en distintas PCR rebajaría el tiempo de genotipado y sería posible la estandarización entre laboratorios.

Pese a todas las desventajas que puedan presentar estos marcadores como el desconocimiento de su secuencia, su modelo de mutación y el alto tiempo y complejidad de genotipado, las ventajas de su uso son mucho más importantes como para obviar su uso en un cromosoma que presenta grandes complicaciones para su estudio. Estos marcadores han aportado una información hasta el momento desconocida. Confirman por la vía macho que hay dos domesticaciones de ganado bovino y aportan evidencias de una tercera que podría haber sucedido en África. Permiten ver que el haplogrupo Y1 es el que presenta mayor diversidad. Han permitido ahondar en la distribución filogeográfica encontrando una familia dentro del haplogrupo Y2 en África que podría ser equivalente al haplogrupo T1 del ADN mitocondrial. Permiten el cálculo de los tiempos de divergencia entre haplogrupos, obteniéndose valores que concuerdan con los datos obtenidos a partir del análisis del ADN mitocondrial. Pero la ventaja más importante es la de añadir una herramienta para estudiar la variabilidad en un cromosoma que presenta baja variabilidad y pequeño tamaño.

Pese a todo el conocimiento generado en esta memoria hay un amplio rango de trabajos a desarrollar. Un cometido futuro en el campo del estudio de las líneas paternas en bovinos es la búsqueda de nuevos SNP así como la búsqueda de genes Y específicos con los que se podría aportar más información en lo referente a la domesticación y el manejo que fueron objeto esos animales. Es necesario encontrar nuevos SNP que nos permitan confirmar hipótesis como una posible domesticación de *B. taurus* en África. Testar nuevos marcadores como marcadores microsatélite y otros IMM no utilizados en esta memoria, así como incrementar el muestreo con

nuevos individuos, sobre todo animales procedentes del Creciente Fértil, dará una visión más acertada de lo que sucedió hace unos 10000 años cuando se inició el proceso de domesticación del *B. primigenius*.





*Conclusiones*



**First:**

Due to the nature of Y-chromosome, a protocol to carefully select Y-specific microsatellites has been developed. It includes tests to ensure amplification only in males, polymorphism, reproducible pattern, a correct paternal compatibility (are inherited as a clone from father to son) and present a repeatable scoring.

**Second:**

IMM markers are a useful tool for studying the phylogeny of the male-specific lineage of the Y chromosome. IMM markers provide more information than the joint use of Y-specific SNPs and microsatellites. As IMM markers have been described in many species, they arise as interesting tools for carrying out within-and between-species evolutionary studies.

**Third:**

The three Y- haplogroups found in bovine and the African Y2 subfamily show a different evolutionary history with multiple paternal origins.

**Fourth:**

The existence of an African Y2 subfamily supports the hypothesis of a domestication of *B. taurus* in Africa.

**Fifth:**

As a summary it can be concluded that the results obtained using Y-specific microsatellites coincide with the information obtained using mitochondrial DNA.

This fact suggests a common expansion of both males and females from the domestication centres. Local introgression events from *B. primigenius* into protodomesticated herds may suggest geographical differences in cattle management.



*Resúmenes*



## 6.1 Summary

Y-chromosome studies may complete surveys using mitochondrial DNA for inferring sex-specific population genetic processes. In this document, genetic techniques were applied to cattle DNA to ascertain questions regarding the origin of the paternal lineages and male phylogeography of *B. taurus* and *B. indicus*. Using single nucleotide polymorphism (SNPs), a total of three lineages have been found for the cattle Y-chromosome. The lineages found in *B. taurus* were Y1 and Y2, meanwhile only one lineage, Y3, was found in *B. indicus*. The aim of this work was to clarify the existence of these lineages and the tool used for resolving that question was the Y-specific microsatellites. A total of 44 microsatellites were tested in both taurine and zebu samples for male specificity, paternal compatibility and scoring repeatability to ascertain their performance for cattle Y-chromosome variability assessment. Y-specific microsatellites added to the Y-specific SNPs previously described increased analytical resolution and allowed at least two different Y2-haplotypic subfamilies to be distinguished, one of them restricted to the African continent and the other one was found in Europe and Asia. New tools, the so called interspersed multilocus microsatellites (IMMs), were developed to allow an improvement of the resolution of bovine Y-chromosome phylogenetic studies. These markers yield several amplified bands of different sizes using a single primer pair in a single PCR reaction from a single male DNA sample. Two out of five IMMs tested for amplification only in male DNA, polymorphism (presence or absence of an amplified band), paternal compatibility and correct and repeatable scoring were used for this study. The use of the IMMs confirmed the findings already described by SNPs but also showed increased resolution with respect to other markers. Its use allowed us to identify a subfamily within the Y2 haplogroup in cattle Y chromosomes sampled from northern Italy, northern Atlantic Europe, Mongolia and Japan. This new subgroup can reflect local introgression events from *B. primigenius* into domesticated cattle in a similar manner than that reflected by the mitochondrial haplogroup Q identified in

mitochondrial DNA. Also, IMMs confirmed the existence of a Y2-haplotypic subfamily restricted to the African continent.

Altogether, these studies have further suggested 1) the introgression of wild sire auroch genetics into domesticated herds, 2) a putative cattle domestication in Africa could included local Y2 wild sires, and 3) the high genetic similarity found in Asian Zebu supports a single domestication event. The lack of samples from the Near East (the only definite centre of taurine cattle domestication) did not allow to check if haplogroup Y1 was originated from a local European post-domestication recruitment of *B. primigenius* sires or was present in the genetic pool domesticated. Overall, the need of further sampling, development and characterization of additional Y-chromosome markers to refine the few phylogenetic studies based on the male lineage has been highlighted.

Interestingly enough, the results obtained with IMMs are highly consistent with the genetic history of the species drawn using mitochondrial DNA. This is surprising because in other species, such as sheep or goat, the genetic information on the male path reflects a different scenario than that given by the mitochondrial DNA. The lack of coincidence between the genetic information from the sire and dam paths is generally explained assuming that the domestication, and the subsequent spreading process of either species, was highly different for each sex. However, cattle have the lowest spreading ability of the livestock species. Its spreading after domestication, whatever the origin might be, was probably very slow and always dependent on human management. In these conditions, the genetic signatures on both males and females would follow similar paths.

The nature and genetic mechanisms affecting IMMs still remain unknown. The existence of markers amplifying several loci is closely linked with the repetitive nature of the Y-chromosome that has been hypothesized to be a mechanism for



maintaining Y-specific genes in a non-recombining environment. Although IMM may have 25 different copies or more, one can hypothesise that UMN0103, with only two copies, is an IMM in an early moment of evolution. The results obtained studying this marker contribute to increase our understanding of IMM.

## 6.2 Resumen

El estudio del cromosoma Y junto con el estudio del ADN mitocondrial permite inferir los procesos genéticos poblacionales determinados por el sexo. En este trabajo, se han aplicado técnicas genéticas en ADN de ganado bovino para establecer cuestiones sobre el origen de los linajes paternos y la filogeografía masculina de *B. taurus* y *B. indicus*. Se usaron polimorfismos de nucleótido simple (SNP), que permitieron definir un total de 3 linajes en el cromosoma Y bovino. En *B. taurus* se describieron los linajes Y1 e Y2, mientras que en *B. indicus* se definió el linaje Y3.

El objetivo de este trabajo es aclarar la existencia de estos tres linajes y la herramienta elegida para resolver esa cuestión fueron los microsatélites específicos del cromosoma Y. Se testaron un total de 44 marcadores en muestras de cebú y toro para determinar polimorfismo, compatibilidad paterno-filial y genotipado repetitivo con el fin de establecer su uso para la evaluación de la variabilidad del cromosoma Y.

Los microsatélites Y específicos sumados a los SNP previamente descritos incrementaron la resolución analítica, esto ha permitido diferenciar dos subfamilias en el linaje Y2, una de ellas restringida al continente africano, mientras que la otra familia se encuentra tanto en Asia como en Europa.

El uso de nuevas herramientas, los microsatélites *multilocus* o IMM, permiten incrementar la resolución de los estudios de filogenia del cromosoma Y bovino. Estos marcadores amplifican numerosas bandas de distintos tamaños a partir de una sola PCR hecha con una sola pareja de cebadores. Se testaron cinco IMM, de los cuales solo dos cumplimentaban los criterios fijados de polimorfismo (presencia o ausencia de una banda amplificada), compatibilidad paterno-filial, genotipado repetitivo y repetibilidad. El uso de los IMM confirma los hallazgos

descritos por los SNP y además se vio que presentaban una mayor resolución. Su uso permite identificar una subfamilia dentro del linaje Y2 en toros, este grupo lo forman muestras del norte de Italia, norte de Europa, Mongolia y Japón. Este nuevo grupo podría reflejar eventos de introgresión de *B. primigenius* en el ganado doméstico de un modo similar reflejado en el haplogrupo Q identificado en el ADN mitocondrial. Además los IMM confirman la existencia de la subfamilia Y2 restringida al continente africano.

La interpretación conjunta de todos los trabajos presentados en esta Memoria sugieren:

- 1) la introgresión de machos de uro en los rebaños domésticos,
- 2) una presunta domesticación de ganado en África que podría haber incluido animales salvajes Y2,
- 3) la alta similitud genética en el cebú apoya un solo evento de domesticación en *B. indicus*.

La falta de muestras procedentes de Oriente Próximo (el único centro de domesticación taurino descrito hasta el momento) no permite verificar si el linaje Y1 se originó a partir del reclutamiento de individuos de *B. primigenius* después de la domesticación, o si el Y1 estaba ya presente en el *pool* genético domesticado. En conjunto hay una necesidad de ampliar el muestreo así como el desarrollo y caracterización de marcadores Y específicos que permitan refinar los estudios filogenéticos realizados hasta el momento por la vía masculina.

Una cuestión que resulta interesante es el hecho de que los resultados obtenidos con los IMM son consistentes con la historia genética dibujada por los análisis del ADN mitocondrial. Esta cuestión resulta llamativa, ya que en otras especies como oveja o cabra, la información obtenida por la vía masculina muestra un escenario diferente a la obtenida por la vía femenina. La falta de coincidencia

entre información genética entre machos y hembras se explica asumiendo que la domesticación y la posterior dispersión fue diferente para cada uno de los sexos. El ganado vacuno presenta una capacidad de dispersión baja, por lo que su dispersión, después de la domesticación, fue muy lenta y probablemente dependiente del manejo del hombre. En esas condiciones la huella genética tanto en machos como en hembras será similar

### 6.3 Resumo

O estudio do cromosoma Y xunto co estudio do ADN mitocondrial permite inferir os procesos xenéticos poblacionais determinados polo sexo. Neste documento aplicáronse as técnicas xenéticas no ADN do gando bovino para establecer cuestións referentes ó orixe dos liñaxes paternos e a filoxeografía masculina de *B. taurus* e *B. indicus*. Usando polimorfismos de nucleotido simple (SNP), atopáronse un total de 3 liñaxes no cromosoma Y do gando bovino. En *B. taurus* describíronse as liñaxes Y1 e Y2, mentras que no *B. indicus* definiuse a liñaxe Y3.

O obxectivo do presente traballo é aclarar a existencia destes tres liñaxes e a ferramenta escollida para resolver esta cuestión foron os microsátélites específicos do cromosoma Y. Testáronse un total de 44 marcadores en mostras de cebú e touro para detectar polimorfismo, compatibilidade paterno-filial e xenotipado repetitivo coa finalidade de avaliar a variabilidade do cromosoma Y.

Os microsátélites Y específicos sumados os SNP previamente descritos, incrementaron a resolución analítica permitindo diferenciar dúas subfamilias dentro da liñaxe Y2, unha delas restrinxida ó continente africano, mentras que a outra atopárase en Asia e Europa. O uso de novas ferramentas, os microsátélites multilocus ou IMM, permiten incrementar a resolución dos estudos filoxenéticos do cromosoma Y bovino. Estes marcadores amplifican numerosas bandas a partir dunha soa PCR feita con unha/cunha única parella de cebadores.

Testaróanse cinco IMM, dos cales so dous cumprimentaban os criterios fixados de polimorfismo (presencia ou ausencia de unha banda amplificada), compatibilidade paterno-filial, xenotipado repetitivo e repetibilidade. O uso dos IMM confirma os achádegos descritos polos SNP e ademáis incrementan a súa resolución. O seu uso permite identificar unha familia dentro do liñaxe Y2 en

touros, este grupo estaría formado por mostras do norte de Italia, norte de Europa, Mongolia e Xapón. Este novo grupo podría reflexar eventos de introgresión de *B. primigenius* no gando doméstico dun modo similar o que se reflexa no haplogrupo Q identificado no ADN mitocondrial. Ademais os IMM confirman a existencia da subfamilia Y restrinxida ó continente africano.

Conxuntamente todos estes traballos suxiren: 1) a introgresión de machos de uro nos rebaños domésticos, 2) unha presunta domesticación de gando en África que podería incluír animais salvaxes Y2, e 3) a alta similitude xenética no cebú apoia un só evento de domesticación en *B.indicus*. A falta de mostras do Oriente Próximo (o único centro de domesticación taurino definido polo momento) non nos permite verificar si a liñaxe Y1 estaba xa presente no *pool* xenético domesticado ou se orixinou despois da domesticación polo reclutamento de individuos de *B. primigenius*. Polo tanto existe unha necesidade de ampliar a mostraxe, así como o desenvolvemento e caracterización de marcadores Y específicos que permitan refinar os estudos filoxenéticos realizados ata o de agora pola vía masculina.

Unha cuestión interesante é o feito de que os resultados obtidos cos IMM son consistentes coa historia xenética dibuxada polos análises de ADN mitocondrial. Esta cuestión resulta chamativa xa que noutras especies, como cabra ou ovella, a información obtida pola vía masculina amosa un escenario diferente que a obtida para a vía feminina. A falta de coincidencia entre información xenética masculina e feminina explícase asumindo que a domesticación e a posterior expansión foi diferente para ambos sexos. O gando vacuno ten pouca capacidade de expansión, polo que a súa dispersión, despois da domesticación, foi moi lenta e probablemente dependente do manexo do home. Nas condicións a pegada xenética tanto en machos como en femias seguiría un camiño similar.



## *Bibliografía*





- Achilli, A., A. Olivieri, M. Pellecchia, C. Uboldi., L. Colli, N. Al-Zahery, M. Accetturo, M. Pala, B.H. Kashani, U.A. Perego, V. Battaglia, S. Fornarino, J. Kalamati, M. Houshmand, R. Negrini, O. Semino, M. Richards, V. Macaulay, L. Ferretti, H.J. Bandelt, P. Ajmone-Marsan, A. Torroni. 2008. Mitochondrial genomes of extinct aurochs survive in domestic cattle. *Curr. Biol.* 18: R157-R158.
- Achilli, A., S. Bonfiglio, A. Olivieri, A. Malusà, M. Pala, B.H. Kashani, U.A. Perego, P. Ajmone-Marsan, L. Liotta, O. Semino, H.J. Bandelt, L. Ferretti, A. Torroni. 2009. The Multifaceted Origin of Taurine Cattle Reflected by the Mitochondrial Genome. *PLoS ONE* 4(6): e5753 doi:10.1371/journal.pone.0005753
- Aikio, P. 1989. The Changing Role of the Reindeer in the Life of the Sámi." In *The Walking Larder: Patterns of Domestication, Pastoralism, and Predation*, edited by Juliet Clutton-Brock, pp.169-184. London: Unwin Hyman.
- Ali, S., S. E. Hasnain. 2002. Molecular dissection of the human Y-chromosome. *Gene* 283: 1-10.
- Baccus, R., N. Ryman, M.H. Smith, D. Reuterwall, D. Cameron. 1983. Genetic variability and differentiation of large grazing mammals. *Jornual of Mammalian*, 64: 109-120.
- Bailey, J.F, M.B. Richards, V.A Macaulay, I.B. Colson, I.T. James, D.G. Bradley, R.E. Hedges, B.C. Sykes. Ancient DNA suggests a recent expansion of European cattle from a diverse wild progenitor species. *Proc. R. Soc. B.* 1996; 263:1467-1473.
- Baig, M., A. Beja-Pereira, R. Mohammad, K. Kulkarni, S. Farah, G. Luikart. 2005. Phylogeography and origin of Indian domestic cattle. *Current Science* 89, 38-40.
- Balaresque, P., A. Sibert, E. Heyer, B. Crouau-Roy. 2006. Unbiased interpretation of haplotypes at duplicated microsatellites. *Annals of Human Genetics* 71, 209-219.
- Beja-Pereira, A., D. Caramelli, C. Lalueza-Fox, C. Vernesi, N. Ferrand, A. Casoli, F. Goyache, L.J. Royo, S. Conti, M. Lari, A. Martini, L. Ouragh, A. Magid, A. Atash, A. Zsolnai, P. Boscato, C. Triantaphylidis, K. Ploumi, L. Sineo, F. Mallegni, P. Taberlet, G. Erhardt, L. Sampietro, J. Bertranpetit, G. Barbujani, G. Luikart, G. Bertorelle. 2006. The origin of European cattle: evidence from modern and ancient DNA. *Proc Natl Acad Sci USA* 103, 8113-8118.
- Beja-Pereira, A., P.R. England, N. Ferrand, S. Jordan, A. Bakhiet, M.A. Abdalla, M. Mashkour, J. Jordana, P. Taberlet, G. Luikart. 2004. African origins of the domestic donkey. *Science* 304, 1781-1783.
- Bishop M.D., S.M. Kappes, J.W. Keele, R.T. Stone, S.L.F. Sunden, G.A. Hawkins, S.S. Toldo, R. Fries, M.D. Grosz, J. Yoo, C.W. Beattie. 1994. A genetic linkage map for cattle. *Genetics*, 136, 619-639.
- Bollongino, R., J. Elsner, J.D. Vigne, J. Burger. 2008. Y-SNPs Do Not Indicate Hybridisation between European Aurochs and Domestic Cattle. *PLoS ONE* 3(10): e3418. doi:10.1371/journal.pone.0003418
- Bonin, A., F. Nicole, F. Pompanon, C. Miaud, P. Taberlet. 2007. Population adaptive index: a new method to help measure intraspecific genetic diversity and prioritize populations for conservation. *Conserv Biol* 21:697-708. doi:10.1111/j.1523-1739.2007.00685.x
- Bourque, G., P.A. Pevzner, G. Tesler. 2004. Reconstructing the genomic architecture of ancestral mammals: lessons from human, mouse, and rat genomes. *Genome Res* 2004;14:507-16.
- Bovine HapMap Consortium, R.A. Gibbs, J.F. Taylor, C.P. Van Tassell, W. Barendse, K.A. Eversole, C.A. Gill, R.D. Green, D.L. Hamernik, S.M. Kappes, S. Lien, L.K.

- Matukumalli, J.C. McEwan, L.V. Nazareth, R.D. Schnabel, G.M. Weinstock, D.A. Wheeler, P. Ajmone-Marsan, P.J. Boettcher, A.R. Caetano, J.F. Garcia, O. Hanotte, P. Mariani, L.C. Skow, T.S. Sonstegard, J.L. Williams, B. Diallo, L. Hailemariam, M.L. Martinez, C.A. Morris, L.O. Silva, R.J. Spelman, W. Mulatu, K. Zhao, C.A. Abbey, M. Agaba, F.R. Araujo, R.J. Bunch, J. Burton, C. Gorni, H. Olivier, B.E. Harrison, B. Luff, M.A. Machado, J. Mwakaya, G. Plastow, W. Sim, T. Smith, M.B. Thomas, A. Valentini, P. Williams, J. Womack, J.A. Woolliams, Y. Liu, X. Qin, K.C. Worley, C. Gao, H. Jiang, S.S. Moore, Y. Ren, X.Z. Song, C.D. Bustamante, R.D. Hernandez, D.M. Muzny, S. Patil, A. San Lucas, Q. Fu, M.P. Kent, R. Vega, A. Matukumalli, S. McWilliam, G. Sclep, K. Bryc, J. Choi, H. Gao, J.J. Grefenstette, B. Murdoch, A. Stella, R. Villa-Angulo, M. Wright, J. Aerts, O. Jann, R. Negrini, M.E. Goddard, B.J. Hayes, D.G. Bradley, M. Barbosa da Silva, L.P. Lau, G.E. Liu, D.J. Lynn, F. Panzitta, K.G. Dodds. 2009 Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. *Science*. 24;324(5926):528-32.
- Bowcock, A.M., A. Ruiz-Linares, J. Tomfohrde, E. Minch, J.R. Kidd, L.L. Cavalli-Sforza. 1994. High resolution of human evolutionary trees with polymorphic microsatellites. *Nature* 31;368(6470):455-7.
- Bradley, D.G., R.T. Loftus, P. Cunningham, D.E. MacHugh. 1998. Genetics and domestic cattle origins. *Evol. Anthropol.* 6, 79–86.
- Bradley, D.G., D.E. MacHugh, R.T. Loftus, R.S. Sow, C. H. Hoste, E.P. Cunningham. 1994. Zebu-taurine variation in Y chromosomal DNA: a sensitive assay for genetic introgression in west African trypanotolerant cattle populations. *Anim Genet.* 25(1):7-12.
- Bradley, D.G., D.E. MacHugh, P. Cunningham, R.T. Loftus. 1996. Mitochondrial diversity and the origins of African and European cattle. *Proc Natl Acad Sci USA.* 93(10): 5131-5.
- Bradley, D.G., D.A. Magee. 2006. Genetics and the origins of domestic cattle. In: M.A. Zeder., D.G. Bradley, E. Emshwiller, B.D. Smith, editors. Documenting domestication: new genetic and archaeological paradigms. London: University of California Press. pp. 317–328.
- Bruford, M.W., D.G. Bradley, G. Luikart. 2003. DNA markers reveal the complexity of livestock domestication. *Nat Rev Genet*; 4: 900-910.
- Burger, J., S. Hummel, B. Herrmann, W. Henke. 1999. DNA preservation: A microsatellite-DNA study on ancient skeletal remains. *Electrophoresis* 20: 1722-1728
- Butler, J.M., R. Schoske. 2005. U.S. population data for the multi-copy Y-STR locus DYS464. *J Forensic Sci.* 50(4):975-7
- Cavalli-Sforza, L.L., E. Minch, 1997. Paleolithic and Neolithic lineages in the European mitochondrial gene pool. *Am J Hum Genet* 61: 247-254.
- Cai, X., H. Chen, S. Wang, K. Xue, C. Lei. 2006. Polymorphisms of two Y chromosome microsatellites in Chinese cattle. *Genetics Selection Evolution* 38, 525–34
- Charlesworth, B. 1994. The effect of background selection against deleterious mutations on weakly selected, linked variants. *Genet Res.* 63(3):213-27.
- Charlesworth, B., D. Charlesworth. 2000. The degeneration of Y chromosomes. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 355: 1563–1572.
- Charlesworth, D., B. Charlesworth, G. Marais. 2005. Steps in the evolution of heteromorphic sex chromosomes. *Heredity* 95: 118–128.

- Charlesworth, D., B. Charlesworth, M.T. Morgan. 1995. The pattern of neutral molecular variation under the background selection model. *Genetics* 141:1619–1632.
- Chen S., B.Z. Lin, M. Baig, B. Mitra, R.J. Lopes, A.M. Santos, D.A. Magee, M. Azevedo, P. Tarroso, S. Sasazaki, S. Ostrowski, O. Mahgoub, T.K. Chaudhuri, Y.P. Zhang, V. Costa, L.J. Royo, F. Goyache, G. Luikart, N. Boivin, D.Q. Fuller, H. Mannen, D.G. Bradley, A. Beja-Pereira. 2010. Zebu cattle are an exclusive legacy of the South Asia Neolithic. *Mol. Biol. Evol.* doi:10.1093/molbev/msp213.
- Chessa, B., F. Pereira, F. Arnaud, A. Amorim, F. Goyache, I. Mainland, R.R. Kao, J.M. Pemberton, D. Beraldi, M. Stear, A. Alberti, M. Pittau, M.H. Banabazi, R. Kazwala, Y.-P. Zhang, J.J. Arranz, B.A. Ali, Z. Wang, M. Uzun, M. Dione, I. Olsaker, L.-E. Holm, U. Saarma, S. Ahmad, N. Marzanov, E. Eythorsdottir, M. Holland, P. Ajmone-Marsan, M.W. Bruford, J. Kantanen, T.E. Spencer, M.E. Palmardini. 2009. Revealing the history of sheep domestication using retrovirus integrations. *Science*, 324, 532–536.
- Clutton-Brock, J. 1979. The Mammalian Remains from the Jericho Tell. *Proceedings of the Prehistoric Society* 45:135–158.
- Clutton-Brock, J. 1992. *Horse Power: A History of the Horse and the Donkey in Human Societies*. Cambridge, Mass.: Harvard University Press.
- Clutton-Brock, J. 1999. *A natural history of domesticated animals*. Cambridge university press, Cambridge.
- Daneau, I., A. Houde, J. F. Ethier, J. G. Lussier, D. W. Silversides. 1995. Bovine SRY gene locus: cloning and testicular expression. *Biol Reprod* 52: 591–599.
- Di Meo, G.P., A. Perucatti, S. Floriot, D. Incarnato, R. Rullo, A. Caputi Jambrenghi, L. Ferretti, G. Vonghia, E. Cribiu, A. Eggen, L. Iannuzzi. 2005. Chromosome evolution and improved cytogenetic maps of the Y chromosome in cattle, zebu, river buffalo, sheep and goat. *Chromos. Res.* 13: 349–355.
- Edwards, A., A. Civitello, H.A. Hammond, C.T. Caskey. 1991. DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. *American Journal of Human Genetics* 49(4):746–56.
- Edwards, A., H.A. Hammond, L. Jin, C.T. Caskey, R. Chakraborty. 1992. Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. *Genomics* 12(2):241–53.
- Edwards, C.J., C. Gaillard, D.G. Bradley, D.E. MacHugh. 2000. Y-specific microsatellite polymorphism in a range of bovid species. *Animal Genetics* 31, 127–130.
- Edwards, C.J., J.F. Baird, D.E. MacHugh. 2007a. Taurine and zebu admixture in Near Eastern cattle: a comparison of mitochondrial, autosomal and Y-chromosomal data. *Animal Genetics Anim Genet.* 2007 Oct;38(5):520–4.
- Edwards, C.J., R. Bollongino, A. Scheu, A. Chamberlain, A. Tresset, J.D. Vigne, J.F. Baird, G. Larson, S.Y. Ho, T.H. Heupink, *et al.* 2007b. Mitochondrial DNA analysis shows a Near Eastern Neolithic origin for domestic cattle and no indication of domestication of European aurochs. *Proc. Biol. Sci.* 274, 1377–1385.
- Egan, C.M., S. Sridhar, M. Wigler, I.M. Hall. 2007. Recurrent DNA copy number variation in the laboratory mouse. *Nat Genet.* 39:1384–1389.
- Ellis, N., P.N. Goodfellow. 1989. The mammalian pseudoautosomal region. *Trends Genet.* 5:406–410.
- Ellison, J.W., X. Li, U. Francke, L.J. Shapiro. 1996. Rapid evolution of human

- pseudoautosomal genes and their mouse homologs. *Mamm Genome* 7(1):25-30.
- Epstein, H., I.L. Mason. 1984. Evolution of Domesticated Animals. Cattle. In: Mason I.L. (ed.) Longman, New York, pp. 6-34.
- Evett, I.W., C. Buffery, G. Willott, D. Stoney. 1991. A guide to interpreting single locus profiles of DNA mixtures in forensic cases. *J Forensic Sci Soc.* 31(1):41-7.
- Excoffier L., G. Laval, S. Schneider. 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evol Bioinf Online* 1: 47–50.
- Ezaz, T., R. Stiglec, F. Veyrunes, J.A.M. Graves. 2006. Relationships between vertebrate ZW and XY sex chromosome systems. *Curr. Biol.* 16: 736–743.
- FAO. 2000. World Watch List for Domestic Animal Diversity. 3rd ed. B. Scherf, ed. FAO, Rome, Italy.
- Felius, M. 1995. Encyclopedia of Cattle Breeds. Misset, Doetinchem, The Netherlands.
- Fincham, A.G., Y.Y. Hu, E. Lau, Z. Pavlova, H.C. Slavkin, M.L. Snead. 1990. Isolation and partial characterization of a human amelogenin from a single fetal dentition using HPLC techniques. *Calcif Tissue Int.*47(2):105-11.
- Forbes, S.H., J.T. Hogg, F.C. Buchanan, A.M. Crawford, F.W. Allendorf. 1995. Microsatellite evolution in congeneric mammals: domestic and bighorn sheep. *Molecular Biology and Evolution* 12(6):1106-13
- Forster, P., R. Harding, A. Torroni, H-J. Bandelt. 1996. Origin and Evolution of Native American mtDNA Variation: A reappraisal. *Am. J. Human Genet.* 59: 935-945.
- Forster, P., A. Röhl, P. Lünemann, C. Brinkmann, T. Zerjal, C. Tyler-Smith, B. Brinkmann. 2000. A Short Tandem Repeat-Based Phylogeny for the Human Y Chromosome. *Am. J. Hum. Genet.* 67: 182-196.
- Fraser, J.A., J. Heitman. 2005. Chromosomal sex-determining regions in animals, plants and fungi. *Curr. Opin. Genet.*15:645–651
- Freeman, J.L., G.H. Perry, L. Feuk, R. Redon, S.A. McCarroll, D.M. Altshuler, H. Aburatani, K.W. Jones, C. Tyler-Smith, M.E. Hurles, N.P. Carter, S.W. Scherer, C. Lee. 2006. Copy number variation: new insights in genome diversity. *Genome Res* 16:949-961.
- Fridolfsson, A.K., H. Cheng, N.G. Copeland, N.A. Jenkins, H.C. Liu, T. Raudsepp, T. Woodage, B. Chowdhary, J. Halverson, H. Ellegren. 1998. Evolution of the avian sex chromosomes from an ancestral pair of autosomes. *Proc. Natl. Acad. Sci.* 95: 8147–8152.
- Fries R., A. Ruvinsky. 1999. The genetics of cattle. CABI Publishing, CAB International, pp. 1-33.
- Frisch J.E., R. Drinkwater, B. Harrison, S. Johnson. 1997. Classification of the southern African sanga and east African shorthorned zebu. *Anim Genet.* 28(2):77-83.
- Gallagher D.S. Jr., J.E. Womack. 1992. Chromosome conservation in the Bovidae. *J. Hered.* 83(4):287-98.
- Gallagher, D. S. Jr., S. K. Davis, M. De Donato, J. D. Burzlaff, J. E. Womack et al. 1999. A molecular cytogenetic analysis of the tribe Bovini (Artiodactyla: Bovidae: Bovinae) with an emphasis on sex chromosome morphology and NOR distribution. *Chromosome Res* 7: 481-492
- Gibson, C., E. Golub, R. Herold, M. Risser, W. Ding et al. 1991. Structure and expression of the bovine amelogenin gene. *Biochemistry* 30: 1075-1079.

- Ginja, C., L. Telo da Gama, M.C.T. Penedo. 2009. Y Chromosome haplotype analysis in Portuguese cattle breeds using SNPs and STRs. *Journal of Heredity* 100, 148–157.
- Giovambattista, G., M.V. Ripoli, J.C. De Luca, P.M. Mirol, J.P. Lirón, F.N. Dulout. 2000. Male-mediated introgression of *Bos indicus* genes into Argentine and Bolivian Creole cattle breeds. *Animal Genetics* 31, 302–305.
- Goldammer, T., R. M. Brunner, M. Schwerin. 1997. Comparative analysis of Y chromosome structure in *Bos taurus* and *B. indicus* by FISH using region-specific, microdissected, and locus-specific DNA probes. *Cytogenet Cell Genet* 77: 238–241.
- Goldstein, D.B., A.R. Linares, L.L. Cavalli-Sforza, M.W. Feldman. 1995. An evaluation of genetic distances for use with microsatellite loci. *Genetics*, 139: 463–471.
- Goloubinoff, P., S. Pääbo, A.C. Wilson. 1993. Evolution of maize inferred from sequence diversity of an *Adh2* gene segment from archaeological specimens. *Proc Natl Acad Sci U.S.A.* 90(5): 1997–2001.
- Götherström, A., C. Anderung, L. Hellborg, R. Elburg, C. Smith, D.G. Bradley, H. Ellegren. 2005. Cattle domestication in the Near East was followed by hybridization with aurochs bulls in Europe. *Proc. R. Soc. B* 272, 2345–2350.
- Graves, J. A. 1995. The origin and function of the mammalian Y chromosome and Y-borne genes – an evolving understanding; *Bioessays* 17 311–320.
- Graves, J.A.M., S. Shetty. 2001. Sex from W to Z—Evolution of vertebrate sex chromosomes and sex determining genes. *J. Exp. Zool.* 290: 449–462.
- Graves, J. A., 2002 Evolution of the testis-determining gene—the rise and fall of SRY. *Novartis Found Symp* 244: 86–97; discussion 97–101, 203–106, 253–107.
- Graves, J.A.M. 2006. Sex chromosome specialization and degeneration in mammals. *Cell.* 124:901–914.
- Grigson, C. 1980. The craniology and relationships of four species of *Bos*. *L., J. Arch. Sci.* 7:3–32.
- Grützner, F., W. Rens, E. Tsend-Ayush, N. El-Mogharbel, P.C.M. O'Brien, R.C. Jones, M.A. Ferguson-Smith, J.A.M. Graves. 2004. In the platypus a meiotic chain of ten sex chromosomes shares genes with the bird Z and mammal X chromosomes. *Nature* 432: 913–917.
- Hagelberg, E., B. Sykes, R. Hedges. 1989. Ancient bone DNA amplified. *Nature.* 342(6249): 485.
- Hagelberg, E., M.G. Thomas, C.E. Cook, A.V. Sher, G.F. Baryshnikov, A.M. Lister. 1994. DNA from ancient mammoth bones. *Nature.* 370(6488): 333–334.
- Hammer, M. F. 1995. A recent common ancestry for human Y chromosomes. *Nature* 378, 376–378.
- Handt, O., M. Höss, M. Krings, S. Pääbo. 1994. Ancient DNA: methodological challenges. *Experientia.* 50: 524–529.
- Hänni, C., V. Laudet, M. Sakka, A. Begue, D. Stehelin. 1990. Amplification of mitochondrial DNA fragments from ancient human teeth and bones. *C R Acad Sci III.* 310(9): 365–370.
- Hanotte, O., C.L. Tawah, D.G. Bradley, M. Okomo, Y. Verjee, J. Ochieng, J.E. Rege. 2000. Geographic distribution and frequency of a taurine *Bos taurine* and an indicine *Bos indicus* Y specific allele amongst sub-Saharan African cattle breeds. *Molecular Ecology* 9, 387–396.
- Hanotte, O., D.G. Bradley, J.W. Ochieng, Y. Verjee, E.W. Hill, J.E. Rege. 2002. African pastoralism: genetic imprints of origins and migrations. *Science* 296, 336–339.

- Hart, B.L. 1985. The Behavior of Domestic Animals. New York: W.H. Freeman and Co. 390pp.
- Hauswirth, W.W., C.D. Dickel, D.A. Lawlor. 1994. DNA analysis of the Windover population. En: Herrmann B., Hummel S. (Eds.). 1994. Ancient DNA. B. Springer Verlag, New York. 104-121.
- Hellborg, L., H. Ellegren. 2003. Y chromosome conserved anchored tagged sequences (YCATS) for the analysis of mammalian male-specific fragments. Mol. Ecol. 12:283-291.
- Hellborg, L., H. Ellegren. 2004. Low levels of nucleotide diversity in mammalian Y chromosomes. Mol Biol Evol. 21(1):158-63.
- Helmer, D., L. Gourichon, H. Monchot, J. Peters, M. Saña Seguí. 2005. The first steps of animal domestication: new archaeozoological approaches. J.-D. Vigne, D. Helmer, J. Peters, Eds. (Oxford: Oxbow Books), pp. 86-95.
- Human Genome Sequencing Consortium .2001. Initial sequencing and analysis of the human genome. Nature 409, 860-921.
- Hurles, M. E., M. A. Jobling. 2001. Haploid chromosomes in molecular ecology: lessons from the human Y. Mol Ecol 10: 1599-1613.
- Iannuzzi, L., G. P. Di Meo, A. Perucatti, A. Eggen, D. Incarnato *et al.* 2001a. A pericentric inversion in the cattle Y chromosome. Cytogenet Cell Genet 94: 202-205.
- Iannuzzi, L., L. Molteni, G. P. Di Meo, A. De Giovanni, A. Perucatti *et al.* 2001b. A case of azoospermia in a bull carrying a Y-autosome reciprocal translocation. Cytogenet Cell Genet 95: 225-227.
- Ihara, N., A. Takasuga, K. Mizoshita, H. Takeda, M. Sugimoto, Y. Mizoguchi, T. Hirano, T. Itoh, T. Watanabe, K.M. Reed, W.M. Snelling, S.M. Kappes, C.W. Beattie, G.L. Bennett, Y. Sugimoto. 2004. A comprehensive genetic map of the cattle genome based on 3802 microsatellites. Genome Res. 2004 14: 1987-1998.
- Jakubiczka, S., F. Schnieders, J. Schmidtke. 1993. A bovine homologue of the human TSPY gene. Genomics 17: 732-735
- Jansen T., P. Forster, M.A. Levine, H. Oelke, M. Hurles, C. Renfrew, J. Weber, K. Olek. 2002. Mitochondrial DNA and the origins of the domestic horse. Proc Natl Acad Sci U S A. 16:10905-10.
- Jegalian K., B. T. Lahn .2001. Why the Y is so weird? Scientific American, Inc. 284: 56-61
- Jobling, M. A., C. Tyler-Smith. 2000. New uses for new haplotypes the human Y chromosome, disease and selection. Trends Genet 16: 356-362.
- Jobling, M.A., C. Tyler-Smith. 2003. The human Y chromosome: an evolutionary marker comes of age. Nat Reviews Genet 4:598-612.
- Just, W., W. Rau, W. Vogel, M. Akhverdian, K. Fredga, J.A. Graves, E. Lyapunova. 1995. Absence of Sry in species of the vole *Ellobius*. Nat Genet. 11(2):117-8.
- Kadwell, M., M. Fernandez, H.F. Stanley, R. Baldi, J.C. Wheeler, R. Rosadio, M.W. Bruford. 2001. Genetic analysis reveals the wild ancestors of the llama and the alpaca. Proc Biol Sci; 268: 2575-2584.
- Kantanen, J., C.J. Edwards, D.G. Bradley, H. Viinalass, S. Thessler, Z. Ivanova, T. Kiselyova, M. Cinkulov, R. Popov, S. Stojanović, I. Ammosov, J. Vilki. 2009. Maternal and paternal genealogy of Eurasian taurine cattle (*Bos taurus*). Heredity. 103(5):404-15.
- Kappes, S.M., J.W. Keele, R.T. Stone, R.A. McGraw, T.S. Sonstegard, T.P. Smith, N.L. Lopez-Corrales, C.W. Beattie. 1997.

- A second-generation linkage map of the bovine genome. *Genome Research* 7, 235-49.
- Kato, Y., S. Sato, X. Cui, Y. Itagaki, S. Sutou. 1995. Cloning and characterization of Bovine *Sry* gene, *Anim. Sci. Tech.* 66: 994-1000.
- Kayser, M., L. Roewer, M. Hedman, L. Henke, J. Henke, S. Brauer, C. Krüger, M. Krawczak, M. Nagy, T. Dobosz, R. Szibor, P. de Knijff, M. Stoneking, A. Sajantila. 2000. Characteristics and frequency of germline mutations at microsatellite loci from the human Y chromosome, as revealed by direct observation in father/son pairs. *Am J Hum Genet.* 5:1580-8.
- Kayser, M., A. Caglià, D. Corach, N. Fretwell, C. Gehrig, G. Graziosi, F. Heidorn, S. Herrmann, B. Herzog, M. Hidding, K. Honda, M. Jobling, M. Krawczak, K. Leim, S. Meuser, E. Meyer, W. Oesterreich, A. Pandya, W. Parson, A. Piccinini, A. Perez-Lezaun, M. Prinz, C. Schmitt, P.M. Schneider, R. Szibor, J. Teifel-Greding, G. Weichhold, P. de Knijff, L. Roewer. 1997. Evaluation of Y chromosomal STRs: A Multicenter Study. *International Journal of Legal Medicine*, 110 (3): 125-133.
- Kayser, M., R. Kittler, A. Erler, M. Hedman, A.C. Lee, A. Mohyuddin, S.Q. Mehdi, Z. Rosser, M. Stoneking, M.A. Jobling, A. Sajantila, C. Tyler-Smith. 2004. A comprehensive survey of human Y-chromosomal microsatellites. *Am J Hum Genet.* 74(6):1183-97.
- Kim, K.S., J.S. Yeo, C.B. Choi. 2002. Genetic diversity of north-east Asian cattle based on microsatellite data. *Anim. Genet.*, 33, 201-204.
- Kimura, M., J.F. Crow. 1964. The number of alleles that can be maintained in a finite population. *Genetics* 49:725-738.
- Kimura, M., T. Ohta. 1978. Stepwise mutation model and distribution of allelic frequencies in a finite population. *Proc Natl Acad Sci USA* 75: 2868-72.
- Kirsch, S., B. Weiss, M. De Rosa, T. Ogata, G. Lombardi, G.A. Rappold. 2000. FISH deletion mapping defines a single location for the Y chromosome stature gene, GCY. *J Med Genet.* 37(8):593-9.
- Kirsch, S., B. Weiss, K. Zumbach, G. Rappold. 2004. Molecular and evolutionary analysis of the growth-controlling region on the human Y chromosome. *Hum Genet.* 114(2):173-81.
- Kohn, M., H. Kehrer-Sawatzki, W. Vogel, J.A.M. Graves, H. Hameister. 2004. Wide genome comparisons reveal the origins of the human X chromosome. *Trends Genet.* 20: 598-603.
- Köhler-Rollefson, I. 1996. The One-Humped Camel in Asia: Origin, Utilization and Mechanisms of Dispersal. In *The Origins and Spread of Agriculture and Pastoralism in Eurasia*, edited by David R. Harris, pp. 282-294. Washington, D.C.: Smithsonian Institution Press
- Kuroda-Kawaguchi, T., H. Skaletsky, L.G. Brown, P.J. Minx, H.S. Cordum, R.H. Waterston, R.K. Wilson, S. Silber, R. Oates, S. Rozen, D.C. Page. 2001. The AZFc region of the Y chromosome features massive palindromes and uniform recurrent deletions in infertile men. *Nature Genet.* 29, 279-286.
- Lahn, B.T., D.C Page. 1997. Functional coherence of the human Y chromosome. *Science.* 278:675.
- Lahn, B.T., D.C Page. 1999. Retroposition of autosomal mRNA yielded testis-specific gene family on human Y chromosome. *Nature Genetics* 21, 429-33.
- Lai, S. J., Y. P. Liu, Y. X. Liu, X. W. Li, Y. G. Yao. 2006. Genetic diversity and origin of Chinese cattle revealed by mtDNA D-loop sequence variation. *Mol Phyl Evol* 38, 146.

- Larson G., K. Dobney, U. Albarella, M. Fang, E. Matisoo-Smith, J. Robins, S. Lowden, H. Finlayson, T. Brand, E. Willerslev, P. Rowley-Conwy, L. Andersson, A. Cooper. 2005. Worldwide phylogeography of wild boar reveals multiple centers of pig domestication. *Science*, 307: 1618-1621.
- Larson G., U. Albarella, K. Dobney, P. Rowley-Conwy, J. Schibler, A. Tresset, J.D. Vigne, C.J. Edwards, A. Schlumbaum, A. Dinu, A. Balăşescu, G. Dolman, A. Tagliacozzo, N. Manaseryan, P. Miracle, L. Van Wijngaarden-Bakker, M. Masseti, D.G. Bradley, A. Cooper. 2007. Ancient DNA, pig domestication, and the spread of the Neolithic into Europe. *Proc Natl Acad Sci U S A*. 104(39):15276-81.
- Legge, T. 1996. The Beginning of Caprine Domestication in Southwest Asia." In *The Origins and Spread of Agriculture and Pastoralism in Eurasia*, edited by David R. Harris, pp. 238–262. Washington, D.C.: Smithsonian Institution Press.
- Li M.H., M. Zerabruk, O. Vangen, I. Olsaker, J. Kantanen. 2007. Reduced genetic structure of north Ethiopian cattle revealed by Y-chromosome analysis. *Heredity* 98, 214-218.
- Liu, W., P. Mariani, C.W. Beattie, L.J. Alexander, F.A. Ponce de León. 2002. A radiation hybrid map for the bovine Y-chromosome. *Mamm. Genome* 13:320-326.
- Liu, W-S., C.W. Beattie, F.A. Ponce de León. 2003. Bovine Y chromosome microsatellite polymorphism. *Cytogenetic and Genome Research* 10, 53-58.
- Liu, W-S., F.A. Ponce de León. 2007a. Mapping of the Bovine Y Chromosome. *Electronic Journal of Biology*, Vol. 3(1): 5-12
- Liu, W.-S., A. Wang, Y. Uno, D. Galitz, C.W. Beattie, F.A. Ponce de León. 2007b. Genomic structure and transcript variants of the bovine DAZL gene *Cytogenet Genome Res.* 116:65-71.
- Loftus, R.T., D.E. MacHugh, D.G. Bradley, P.M. Sharp, E.P. Cunningham. 1994a. Mitochondrial genetic variation in European, African and Indian cattle populations. *Animal Genetics*, 25: 265-27.
- Loftus, R.T., D.E. MacHugh, D.G. Bradley, P.M. Sharp, E.P. Cunningham. 1994b. Evidence for two independent domestications of cattle. *Proc Natl Acad Sci U S A*. 91: 2757-2761.
- Luikart, G., L. Gielly, L. Excoffier, J.D. Vigne, J. Bouvet, P. Taberlet. 2001. Multiple maternal origins and weak phylogeographic structure in domestic goats. *Proc Natl Acad Sci U S A*. 10:5927-32.
- Luo, S.J., W.E. Johnson., V.A. David, M. Menotti-Raymond, R. Stanyon, Q.X. Cai, T. Beck, N. Yuhki, J. Pecon-Slaterry, J.L. Smith, S.J. O'Brien. 2007. Development of Y chromosome intraspecific polymorphic markers in the Felidae. *J Hered.* 98(5):400-13.
- MacHugh, D. E., M. D. Shriver, R. T. Loftus, P. Cunningham, D. G. Bradley. 1997. Microsatellite DNA variation and the evolution, domestication and phylogeography of taurine and zebu cattle (*Bos taurus* and *Bos indicus*). *Genetics*, 146: 1071-1086.
- Magee, D.A., C. Meghen, S. Harrison, C.S. Troy, T. Cymbron, C. Gaillard, A. Morrow, J.C. Maillard, D.G. Bradley. 2002. A partial african ancestry for the creole cattle populations of the Caribbean. *J Hered* 93: 429–432.
- Manceau, V., L. Despres, J. Bouvet, P. Taberlet. 1999. Systematics of the genus *Capra* inferred from mitochondrial DNA sequence data. *Mol Phylogenet Evol*; 13: 504-510.
- Mannen, H., Y. Nagata, S. Tsuji. 2001. Mitochondrial DNA reveal that



- domestic goat (*Capra hircus*) are genetically affected by two subspecies of bezoar (*Capra aegagurus*). *Biochem Genet*; 39: 145-154.
- Mannen, H., M. Kohno, Y. Nagata, S. Tsuji, D.G. Bradley, J.S. Yeo, D. Nyamsamba, Y. Zagdsuren, M. Yokohama, K. Nomura, *et al.* 2004. Independent mitochondrial origin and historical genetic differentiation in North Eastern Asian cattle. *Mol. Phylogenet. Evol.* 32, 539-544.
- Manwell, C., C.M.A. Baker. 1980a. Chemical classification of cattle I. Breed groups. *Animal Blood Groups and Biochemical Genetics*, 11: 127-150.
- Manwell, C., C.M.A. Baker. 1980b. Chemical classification of cattle II. Phylogenetic tree and specific status of the zebu. *Animal Blood Groups and Biochemical Genetics*, 3: 295-320.
- Marshall, F. 1989. Rethinking the role of *Bos indicus* in sub-Saharan Africa. *Current Anthropology* 30: 235-240.
- Mathias, N., M. Bayes, C. Tyler-Smith. 1994. Highly informative compound haplotypes for the human Y chromosome. *Human Molecular Genetics* 3, 115-23.
- Matsubara, K., H. Tarui, M. Toriba, K. Yamada, C. Nishida-Umehara, K. Agata, Y. Matsuda. 2006. Evidence for different origin of sex chromosomes in snakes, birds, and mammals and step-wise differentiation of snake sex chromosomes. *Proc. Natl. Acad. Sci.* 103: 18190-18195.
- Matthews, M.E., K.C. Reed. 1991. A DNA sequence that is present in both sexes of Artiodactyla is repeated on the Y chromosome of cattle, sheep and goats. *Cytogenetics and Cell Genetics* 56: 40-44.
- Meadow, H.R. 1993. Animal domestication in the middle east: a revised view from the eastern margin. In Possehl, G.L. (ed.), *Harappan civilization a recent perspective*, Oxford & IBH, pp. 295-320.
- Meadows, J.R.S., O. Hanotte, C. Drögemüller, J. Calvo, R. Godfrey, D. Coltman, J.F. Maddox, N. Marzanov, J. Kantanen, J.W. Kijas. 2006. Globally dispersed Y chromosomal haplotypes in wild and domestic sheep. *Animal Genetics* 37, 444-453.
- Meadows, J.R., I. Cemal, O. Karaca, E. Gootwine, J.W. Kijas. 2007. Five ovine mitochondrial lineages identified from sheep breeds of the near East. *Genetics*. 3:1371-9.
- Mertens, A. 1906. Der Ur, Bos primigenius Bojanus, mit besonderer Berücksichtigung der im städtischen "Museum für Natur- und Heimatkunde" zu Magdeburg befindlichen Reste.-Abhandlungen und Berichte des Museums für Natur- und Heimatkunde zu Magdeburg 1 (2), 1 - 119.
- Mezzelani, A., Y. Zhang, L. Redaelli, B. Castiglioni, P. Leone, J.L. Williams, S.S. Toldo, G. Wigger, R. Fries, L. Ferretti. 1995. Chromosomal localization and molecular characterization of 53 cosmid-derived bovine microsatellites. *Mammalian Genome* 6, 629-635.
- Mumm, S., B. Molini, J. Terrell, A. Srivastava, D. Schlessinger. 1997. Evolutionary features of the 4-Mb Xq21.3 XY homology region revealed by a map at 60-kb resolution. *Genome Res.* 7, 307-314.
- Murphy W.J., A.J.P. Wilkerson, T. Raudsepp, R. Agarwala, A.A. Schaffer, R. Stanyon B.P. Chowdhary. 2006. Novel gene acquisition on carnivore Y chromosomes. *PLoS Genet.* 2: 353-363.
- Naderi, S., H.R. Rezaei, F. Pompanon, M.G. Blum, R. Negrini, H.R. Naghash, O. Balkiz, M. Mashkour, O.E. Gaggiotti, P. Ajmone-Marsan, A. Kence, J.D. Vigne, P. Taberlet. 2008. The goat domestication process inferred from large-scale mitochondrial DNA analysis of wild and domestic individuals. *Proc Natl Acad Sci U S A.* 46:17659-64.

- Naderi, S., H.R. Rezaei, P. Taberlet, S. Zundel, S.A. Rafat, H.R. Naghash, M.A. el-Barody, O. Ertugrul, F. Pompanon, Econogene Consortium. 2007. Large-scale mitochondrial DNA analysis of the domestic goat reveals six haplogroups with high diversity. *PLoS One*. 2(10):e1012.
- Nadir, E., H. Margalit, T. Gallily, S. A. Ben-Sasson. 1996. Microsatellite spreading in the human genome: evolutionary mechanisms and structural implications. *Proc Natl Acad Sci U S A*. 93: 6470-6475.
- Nanda, I., E. Zend-Ajusch, Z. Shan, F. Grützner, M. Scharl, D.W. Burt, M. Koelher, V.M. Fowler, G. Goodwin, W.J. Schneider, S. Mizuno, G. Dechant, T. Haaf, M. Schmid. 2000. Conserved synteny between the chicken Z sex chromosome and human chromosome 9 includes the male regulatory gene DMRT1: A comparative (re)view on avian sex determination. *Cytogenet. Cell Genet.* 89: 67-78.
- Novoa, C., J.C. Wheeler. 1984. Llama and Alpaca. In *Evolution of Domesticated Animals*, edited by Ian Mason, pp. 116-128. London: Longman.
- Odelberg, S.J., D.B. Demers, E.H. Westin, A.A. Hossaini. 1988. Establishing paternity using minisatellite DNA probes when the putative father is unavailable for testing. *J Forensic Sci.* 4:921-8.
- Ogata, T., N. Matsuo. 1993. Sex chromosome aberrations and stature: deduction of the principal factors involved in the determination of adult height. *Hum Genet.* 6:551-62.
- Ohno, S. 1967. Sex chromosomes and sex linked genes. Springer, Berlin.
- Olsen, S. J. 1990. Fossil Ancestry of the Yak: Its Cultural Significance and Domestication in Tibet. *Proceedings of the Philadelphia Academy of Natural Sciences* 142: 73-100.
- Olsen, S. J. 1993. Evidence of Early Domestication of the Water Buffalo in China. In *Skeletons in Her Cupboard: Festschrift for J. Clutton-Brock*, edited by A. Clason, S. Payne, and H-P. Uerpmann, pp. 151-156. Oxford: Oxbow Monograph 34.
- Oota, H., W. Settheetham-Ishida, D. Tiwawech, T. Ishida, M. Stoneking. 2001. Human mtDNA and Y-chromosome variation is correlated with matrilineal versus patrilineal residence. *Nat Genet* 29: 20-21.
- Pääbo, S., H. Poinar, D. Serre, V. Jaenicke-Despres, J. Hebler, N. Rohland, M. Kuch, J. Krause, L. Vigilant, M. Hofreiter. 2004. Genetic analyses from ancient DNA. *Annual Review of Genetics*. 38: 645-679.
- Pääbo, S. 1985a. Preservation of DNA in ancient Egyptian mummies. *J. Arch. Sci.* 12: 411-417.
- Pääbo, S. 1985b. Molecular cloning of ancient Egyptian mummy DNA. *Nature*. 314: 644-645.
- Pääbo, S. 1989. Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification. *Proc Natl Acad Sci U.S.A.* 86(6): 1939-1943.
- Page, D.C., M.E. Harper, J., D. Love. 1984. Botstein. Occurrence of a transposition from the X-chromosome long arm to the Y-chromosome short arm during human evolution. *Nature* 311, 119-123.
- Payne, W.J.A. 1991. *Domestication: a forward step in civilisation*. In: Hickman, C.G. (ed.) *Cattle Genetic Resources*. Wold Animal Science Vol. B7, Elsevier, Amsterdam.
- Perry, G.H., J. Tchinda, S.D. McGrath, J. Zhang, S.R. Picker, A.M. Cáceres, A.J. Iafrate, C. Tyler-Smith, S.W. Scherer, E.E. Eichler, A.C. Stone, C. Lee. 2006. Hotspots for copy number variation in chimpanzees and humans. *Proc Natl Acad Sci USA*, 103:8006-8011.

- Pidancier, N., S. Jordan, G. Luikart, P. Taberlet. 2006. Evolutionary history of the genus *Capra* (Mammalia, Artiodactyla): discordance between mitochondrial DNA and Y-chromosome phylogenies. *Mol. Phyl. Evol.* 40: 739–49.
- Poloumienko, A. 2004. Cloning and comparative analysis of the bovine, porcine, and equine sex chromosome genes ZFX and ZFY. *Genome.* 47:74-83.
- Pompanon, F., A. Bonin, E. Bellemain, P. Taberlet. 2005. Genotyping errors: causes, consequences and solutions. *Nature Reviews Genetics* 6:847–859.
- Price, E. 1984. Behavioral aspects of animal domestication. *Q. Rev. Biol.* 59: 1-32.
- Quinn, A.E., A. Georges, S.D. Sarre, F. Guarino, T. Ezaz, J.A.M. Graves. 2007. Temperature sex reversal implies sex gene dosage in areptile. *Science* 316: 411.
- Rao, E., B. Weiss, M. Fukami, A. Rump, B. Niesler, A. Mertz, K. Muroya, G. Binder, S. Kirsch, M. Winkelmann, G. Nordsiek, U. Heinrich, M.H. Breuning, M.B. Ranke, A. Rosenthal, T. Ogata, G.A. Rappold. 1997. Pseudoautosomal deletions encompassing a novel homeobox gene cause growth failure in idiopathic short stature and Turner syndrome. *Nature Genet.* 16, 54-63.
- Raudsepp, T., E.J. Lee, S.R. Kata, C. Brinkmeyer, J.R. Mickelson, L.C. Skow, J.E. Womack, B.P. Chowdhary. 2004. Exceptional conservation of horse-human gene order on X chromosome revealed by high-resolution radiation hybrid mapping. *Proc Natl Acad Sci USA.* 101:2386–91.
- Raymond, C.S., J.R. Kettlewell, B. Hirsch, V.J. Bardwell, D. Zarkower. 1999. Expression of *Dmrt1* in the genital ridge of mouse and chicken embryos suggests a role in vertebrate sexual development. *Dev. Biol.* 215: 208–220.
- Rhesus Macaque Genome Sequencing and Analysis Consortium, *et al.* 2007. Evolutionary and biomedical insights from the rhesus macaque genome. *Science.* 13;316(5822):222-34.
- Rice, W.R. 1987. Genetic Hitchhiking and the Evolution of Reduced Genetic Activity of the Y Sex Chromosome. *Genetics.* 116(1): 161–167.
- Rogan, P.K., J.J. Salvo. 1990. Study of nucleic acids isolated from ancient remains. *Yearb. Phys. Anthropol.* 33: 195-214.
- Rogers, A.R. 2002. The theoretical mismatch distribution. *Lecture Notes on Gene Genealogies:* 7: 30-46.
- Roldan E.R., M. Gomendio. 1999, The Y chromosome as a battle ground for sexual selection. *Trends in Ecology & Evolution.* 14:2:58-62.
- Ross, M.T., D.V. Grafham, A.J. Coffey, S. Scherer, K. McLay, D. Muzny, M. Platzer, G.R. Howell, C. Burrows, C.P. Bird, A. Frankish, F.L. Lovell, K.L. Howe, J.L. Ashurst, R.S. Fulton, R. Sudbrak, G. Wen, M.C. Jones, M.E. Hurles, T.D. Andrews, C.E., Scott S. Searle, J. Ramser, A. Whittaker, R. Deadman, N.P. Carter, S.E. Hunt, *et al.* 2005 The DNA sequence of the human X chromosome. *Nature.* 434:325–37.
- Rosser, Z.H., T. Zerjal, M.E. Hurles, M. Adojaan, D. Alavantic, A. Amorim, W. Amos, M. Armenteros, E. Arroyo, G. Barbujani, G. Beckman, L. Beckman, J. Bertranpetit, E. Bosch, D.G. Bradley, G. Brede, G. Cooper, H.B. Côte-Real, P. de Knijff, R. Decorte, Y.E. Dubrova, O. Evgrafov, A. Gilissen, S. Glisic, M. Gölge, E.W. Hill, A. Jeziorowska, L. Kalaydjieva, M. Kayser, T. Kivisild, S.A. Kravchenko, A. Krumina, V. Kucinskas, J. Lavinha, L.A. Livshits, P. Malaspina, S. Maria K., McElreavey, T.A. Meitinger, A.V. Mikelsaar, R.J. Mitchell, K. Nafa, J. Nicholson, S. Nørby, A. Pandya, J. Parik, P.C. Patsalis, L. Pereira, B. Peterlin, G.

- Pielberg, M.J. Prata, C. Previderé, L. Roewer, S. Rootsi, D.C. Rubinsztein, J. Saillard, F.R. Santos, G. Stefanescu, B.C. Sykes, A. Tolun, R. Villems, C. Tyler-Smith, M.A. Jobling. 2000. Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. *Am J Hum Genet.* 67(6):1526-43.
- Saifi, G.M., H.S. Chandra. 1999 An apparent excess of sex- and reproductionrelated genes on the human X chromosome. *Proc R Soc Lond B Biol Sci*, 266:203-9.
- Salamini, F., H. Ozkan, A. Brandolini, R. Schäfer-Pregl, W. Martin. 2002. Genetics and geography of wild cereal domestication in the near east. *Nat Rev Genet.* 3(6):429-41.
- Saxena, R., J.W. de Vries, S. Repping, R.K. Alagappan, H. Skaletsky, L.G. Brown, P. Ma, E. Chen, J.M. Hoovers, D.C. Page. 2000. Four DAZ genes in two clusters found in the AZFc region of the human Y chromosome. *Genomics* 67; 256-267.
- Schneider, P.M., S. Meuser, W. Waiyawuth, Y. Seo, C. Rittner. 1998. Tandem repeat structure of the duplicated Y-chromosomal STR locus DYS385 and frequency studies in the German and three Asian populations. *Forensic Sci Int.* 97(1):61-70.
- Schneider, S., D. Roessli, L. Excoffier. 2000. Arlequin ver. 2000: a software for population genetics data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland. URL: <http://lgb.unige.ch/arlequin>.
- Schwartz, A., D. C. Chan, L. G. Brown, R. Alagappan, D. Pettay, C. Disteché, B. McGillivray, A. de la Chapelle, D. Page. 1998. Reconstructing hominid Y evolution: X-homologous block, created by X-Y transposition, was disrupted by Yp inversion through LINE-LINE recombination. *Hum. Mol. Genet.* 7, 1-11.
- Seielstad M. T., E. Minch, L. L. Cavalli-Sforza. 1998. Genetic evidence for a higher female migration rate in humans. *Nat Genet* 20: 278-280
- Skaletsky, H., T. Kuroda-Kawaguchi, P.J. Minx, H.S. Cordum, L. Hillier, L.G. Brown, S. Repping, T. Pyntikova, J. Ali, T. Bieri, A. Chinwalla, A. Delehaanty, K. Delehaanty, H. Du, G. Fewell, L. Fulton, R. Fulton, T. Graves, S.F. Hou, P. Latrielle, S. Leonard, E. Mardis, R. Maupin, J. McPherson, T. Miner, W. Nash, C. Nguyen, P. Ozersky, K. Pepin, S. Rock, T. Rohlfing, K. Scott, B. Schultz, C. Strong, A. Tin-Wollam, S.P. Yang, R.H. Waterston, R.K. Wilson, S. Rozen, D.C. Page. 2003. The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature.* 423:825-37.
- Smith, B.D. 1995. The emergence of agriculture. Scientific American Library New York, New York.
- Sonstegard, T.S., W. Barendse, G.L. Bennett, G.A. Brockmann, S. Davis C., Droegemuller, E. Kalm, S.M. Kappes, C. Kühn, Y. Li, M. Schwerin, J. Taylor, H. Thomsen, C.P. Van Tassell, C.C. Yeh. 2001. Consensus and comprehensive linkage maps of the bovine sex chromosomes. *Animal Genetics.* 32(2):115-117.
- Steinemann, S., M. Steinemann. 2005. Y chromosomes: Born to be destroyed. *Bioessays* 27: 1076-1083.
- Stock, F., C.J. Edwards, R. Bollongino, E.K. Finlay, J. Burger, D.G. Bradley. 2009. Cytochrome b sequences of ancient cattle and wild ox support phylogenetic complexity in the ancient and modern bovine populations *Anim Genet.* 40(5):694-700.
- Stranger, B.E., M.S. Forrest, M. Dunning, C.E. Ingle, C. Beazley, N. Thorne, R. Redon, C.P. Bird, A. de Grassi, C. Lee, C. Tyler-Smith, N. Carter, S.W. Scherer, S. Tavaré, P. Deloukas, M.E. Hurles, E.T. Dermitzakis. 2007. Relative impact of

- nucleotide and copy number variation on gene expression phenotypes. *Science*. 315:848-853.
- Sun, C., H. Skaletsky, B. Birren, K. Devon, Z. Tang, S. Silber, R. Oates, D.C. Page. 1999. An azoospermic man with a de novo point mutation in the Y-chromosomal gene USP9Y. *Nature Genet.* 23, 429-432.
- Sundqvist, A.K., H. Ellegren, M. Olivier, C. Vilà. 2001. Y chromosome haplotyping in Scandinavian wolves (*Canis lupus*) based on microsatellite markers. *Mol. Ecol.* 10(8):1959-66.
- Svensson, E., A. Götherström. 2008. Temporal fluctuations of Y-chromosomal variation in *Bos Taurus*. *Biological Letters* 4, 752-754.
- Takada, T., Y. Kikkawa, H. Yonekawa, S. Kawakami, T. Amano. 1997. Bezoar (*Capra aegagrus*) is a matriarchal candidate for ancestor of domestic goat (*Capra hircus*): evidence from the mitochondrial DNA diversity. *Biochem Genet*; 35: 315-326.
- Tellam, R.L., D.G. Lemay, C.P. Van Tassell, H.A. Lewin, K.C. Worley, C.G. Elsik. 2009. Unlocking the bovine genome. *BMC Genomics*. 24;10:193.
- Thomson, R., J. K. Pritchard, P. Shen, P. J. Oefner, M. W. Feldman. 2000. Recent common ancestry of human Y chromosomes: evidence from DNA sequence data. *Proc. Natl Acad. Sci. USA* 97, 7360-7365.
- Tosi, A.J., J.C. Morales, D.J. Melnick. 2003. Paternal, maternal, and biparental molecular markers provide unique windows onto the evolutionary history of macaque monkeys. *Evolution*. 57(6):1419-35.
- Troy, C.S., D.E. MacHugh, J.F. Bailey, D.A. Magee, R.T. Loftus, P. Cunningham, A.T. Chamberlain, B.C. Sykes, D.G. Bradley. 2001. Genetic evidence for Near- Eastern origins of European cattle. *Nature* 410, 1088-1099.
- Vaiman, D., M. Imam-Ghali, K. Moazami-Goudarzi, G. Guerin, C. Grohs, H. Leveziel, N. Saidi-Mehtar. 1994. Conservation of a syntenic group of microsatellite loci between cattle and sheep. *Mammalian Genome* 5, 310-314.
- Vallender, E.J., B.T. Lahn. 2004. How mammalian sex chromosomes acquired their peculiar gene content. *Bioessays* 26: 159-169.
- Vallender, E.J., B.T. Lahn. 2006. Multiple Independent Origins of Sex Chromosomes in Amniotes *Proc Natl Acad Sci* 103(4): 18031-18032.
- Van Hooft, W.F., A.F. Groen, H.H. Prins. 2002. Phylogeography of the African buffalo based on mitochondrial and Y-chromosomal loci: Pleistocene origin and population expansion of the Cape buffalo subspecies. *Mol Ecol.* 11(2):267-79.
- Van Laere, A-S., W. Coppieters, M. Georges. 2008. Characterization of the bovine pseudoautosomal boundary: Documenting the evolutionary history of mammalian sex chromosomes. *Genome Res.* 18: 1884-1895.
- Van't Hof, A.E., P.M. Brakefield, I.J. Saccheri, B.J. Zwaan. 2007. Evolutionary dynamics of multilocus microsatellite arrangements in the genome of the butterfly *Bicyclus anynana*, with implications for other Lepidoptera. *Heredity* 98, 320-328.
- Van Vuure, C. 2005. Retracing the Aurochs: History, Morphology and Ecology of an Extinct Wild Ox. Pensoft Publishers. Sofia-Moscow.
- Venter, J. C. et al. 2001. The sequence of the human genome. *Science* 291, 1304-1351.
- Verkaar, E.L.C., C. Zijlstra, E.M. Van't Veld, K. Boutaga, D.C.J. Van Bostel, J.A. Lenstra. 2004. Organization and concerted

- evolution of the ampliconic Y-chromosomal TSPY genes from cattle, *Genomics* 84: 468–474.
- Vilà, C., P. Savolainen, J.E. Maldonado, I.R. Amorim, J.E. Rice, R.L. Honeycutt, K.A. Crandall, J. Lundeberg, R.K. Wayne. 1997. Multiple and ancient origins of the domestic dog. *Science*; 276: 1687–1689.
- Vilà C., J.A. Leonard, A. Götherström, S. Marklund, K. Sandberg, K. Liden, R.K. Wayne, H. Ellegren. 2001. Widespread origins of domestic horse lineages. *Science*. 291(5503):474–7.
- Veyrunes F., P.D. Waters, P. Miethke, W. Rens, D. McMillan, A.E. Alsop, F. Grützner, J.E. Deakin, C.M. Whittington, K. Schatzkamer, C.L. Kremitzki, T. Graves, M.A. Ferguson-Smith, W. Warren, J.A.M. Graves. 2008. Bird-like sex chromosomes of platypus imply recent origin of mammal sex chromosomes. *Genome Res.* 18:965–973
- Vogel, T., S. Borgmann, F. Dechend, W. Hecht, J. Schmidtke. 1997a. Conserved Y-chromosomal location of TSPY in Bovidae. *Chromosome Res* 5: 182–185.
- Vogel, T., F. Dechend, E. Manz, C. Jung, S. Jakubiczka, S. Fehr, J. Schmidtke, F. Schnieders. 1997b. Organization and expression of bovine TSPY. *Mamm Genome* 8: 491–496.
- Wallner, B., Piumi, F., Brem, G., Müller, M., Achmann, R. 2004. Isolation of Y chromosome-specific microsatellites in the horse and cross-species amplification in the genus *Equus*. *J Hered.* 95(2):158–64.
- Wang, A., H. Yasue, L. Li, M. Takashima, F.A. Ponce de León, W.-S. Liu. 2008. Molecular characterization of the bovine chromodomain Y-like genes. *Anim Genet.* 2008 Jun;39(3):207–16
- Wendorf, F., R. Schild. 1994. Are the early holocene cattle in the eastern Sahara domestic or wild? *Evol. Anthropol.* 3:118–128.
- Wiegand, P., Schürenkamp, M., Schütte, U. DNA extraction from mixtures of body fluid using mild preferential lysis. *Int J Legal Med.* 1992;104(6):359–60.
- Willerslev, E., A. Cooper. 2005. Ancient DNA. *Proc Biol Sci.* 7;272(1558):3–16.
- Wilson A.F., A.J.M. Sorant. 2000. Equivalence of Single- and Multilocus Markers: Power to Detect Linkage with Composite Markers Derived from Biallelic Loci. *Am. J. Hum. Genet.* 66:1610–1615.
- Wind, A.E., Larkin, D.M., Green, C.A., Elliott, J.S., Olmstead, C.A., Chiu, R., Schein, J.E., Marra, M.A., Womack, J.E., Lewin, H.A. 2005. A high-resolution whole-genome cattle-human comparative map reveals details of mammalian chromosome evolution. *Proc Natl Acad Sci U S A.* 20;102(51):18526–31.
- Wyckoff, G.J., W. Wang, C.-I. Wu. 2000. Rapid Evolution of Male Reproductive Genes in the Descent of Man. *Nature*, 403(6767):304–9.
- Xiao, C., K. Tsuchiya, S. Sutou. 1998. Cloning and mapping of bovine ZFX gene to the long arm of the X-chromosome (Xq34) and homologous mapping of ZFY gene to the distal region of the short arm of the bovine (Yp13), ovine (Yp12-p13), and caprine (Yp12-p13) Y chromosome. *Mamm Genome* 9: 125–130.
- Zeder, M.A., B. Hesse. 2000. The initial domestication of goats (*Capra hircus*) in the Zagros mountains 10,000 years ago. *Science*; 287: 2254–2257.
- Zeder, M. A., D.G., Bradley, E. Emshwiller, B.D., Smith. 2006 Documenting Domestication: New Genetic and Archaeological Paradigms (Univ. California Press, Berkeley), in press.

- Zeuner, F.E. 1963. A History of Domesticated Animals. Harper & Row, New York.
- Zerjal, T., R.S. Wells, N. Yuldasheva, R. Ruzibakiev, C. Tyler-Smith. 2002. A genetic landscape reshaped by recent events: Y-chromosomal insights into central Asia. *Am J Hum Genet.*;71:466–482
- Zhivotovsky, L.A., P.A. Underhill, C. Cinnioğlu, M. Kayser, B. Morar, T. Kivisild, R. Scozzari, F. Cruciani, G. Destro-Bisol, G. Spedini, G.K. Chambers, R.J. Herrera, K.K. Yong, D. Gresham, I. Tournev, M.W. Feldman, L. Kalaydjieva. 2004. The effective mutation rate at Y chromosome short tandem repeats, with application to human population divergence time. *Am. J. Hum. Genet.* **74**: 50–61.







*Anexos*



## Supplementary Material. PCR conditions and DNA sequences of the markers used.

### 1. PCR conditions and genotyping

PCR was performed in a GenAmp PCR system 9700 (Applied Biosystems). The final reaction volume was 10  $\mu$ l. The reaction mixture consisted on 50-100 ng. of genomic DNA, 0.2 units of Taq polymerase (Biotools, Madrid, Spain), and a final concentration of 0.2  $\mu$ M of each primer (Sigma-Aldrich, Madrid, Spain), 200  $\mu$ M of each dNTP (Biotools, Madrid, Spain) and 2 mM  $MgCl_2$ . The PCR protocol included an initial step of 95°C (3 min), followed by 30 cycles of 30 s at 95°C for DNA denaturation, 30 s for primer annealing at 55, 58 and 60 °C and 30 s at 72°C for primer extension. An additional final step of 72°C for 90 min was included to avoid n-1 band presence when using capillary electrophoresis.

Genotyping was carried out using an Automatic Sequencer ABI-310 (Applied Biosystems Inc., Foster City, CA). The size characterization of PCR products was carried out using GenMapper software (Applied Biosystems Inc., Foster City, CA).

### 2- DNA sequence of *INRA124*. Polymorphisms are underlined.

X71546 tgtgtgtatg tgtgtgtgtg tgtgtgttgg tattctcaga cctgtgccct  
ggcgactatg aaacagtttt gtgtgtgtgt gtgtgtgttt

Bull-1 tgtgtgtatg tgtgtgtgtg tgtgtgtttg tattctcaga cctgtgccct  
ggcgactatg aaacagtttt gtgtgtgtgt gtgtgtgttt

Bull-9 tgtgtgtatg tgtgtgtgtg tgtgtgtttg tattctcaga cctgtgccct  
ggcgactatg aaacagtttt gtgtgtgtgt gtgtgtgttt

Cow-C tgtgtgtatg tgtgtgtgtg tgtgtgtttg tattctcaga cctgtgccct  
ggcgactatg aaacagtttt gtgtgtgtgt gtgtgtgttt

Cow -G tgtgtgtatg tgtgtgtgtg tgtgtgtttg tattctcaga cctgtgccct  
ggcgactatg aaacagtttt gtgtgtgtgt gtgtgtgttt

### 3- DNA sequence of *INRA126*. Polymorphisms are underlined.

X71553 gagatgggtg gctggggaac aagagtagaa gacttacttt ccatttcttt  
 aaaacttttag aatttgtccc attcttgtgt gtgtgtgtgt gtgtgtgtac  
 atatatctct tatttagaaa ataaataagt aaataaaatt

Bull-1 gagatgggtg gctggggaac aagagtagaa gacttacttt ccatttcttt  
 aaaacttttag aatttgttcc attcttgtgt gtgtgtgtgt gtgtgtgtac  
 atatgtctct tatttagaaa ataaataagt aaataaaatt

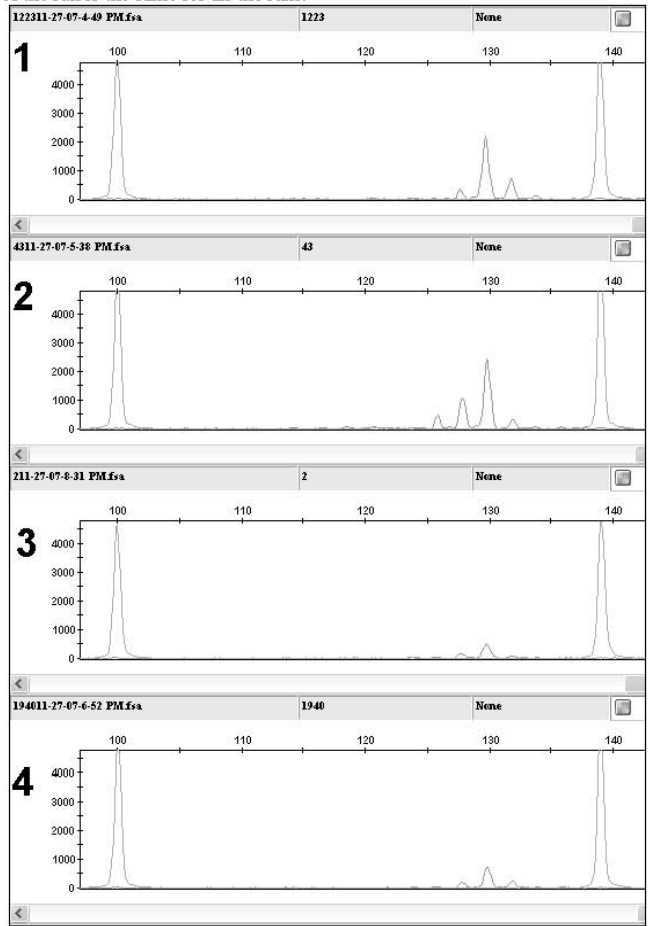
Bull-9 gagatgggtg gctggggaac aagagtagaa gacttacttt ccatttcttt  
 aaaacttttag aatttgttcc attcttgtgt gtgtgtgtgt gtgtgtgtac  
 atatatctct tatttagaaa ataaataagt aaataaaatt

Cow-C gagatgggtg gctggggaag aagagtagaa gacttacttt ccatttcttt  
 aaaacttttag aatttgttcc attcttgtgt gtgtgtgtgt gtgtgtgtac  
 atatgtctct tatttagaaa ataaataagt aaataaaatt

Cow-G gagatgggtg gctggggaac aagagtagaa gacttacttt ccatttcttt  
 aaaacttttag aatttgttcc attcttgtgt gtgtgtgtgt gtgtgtgtac  
 atatatctct tatttagaaa ataaataagt aaataaaatt

**Figure S1.** Genmapper image of *INRA124* amplification (green lines due to VIC labelling of PCR products). Bands of different size form the GenScan 500-LIZ standard (orange).

The first and second plots represent male amplification. The third and fourth plots represent female amplification. All the samples were diluted to the same concentration, amplified at the same time, and with the same PCR mix. All the PCR products were diluted (1/100) before run in an ABB310 sequencer. Fluorescence scale on the left side of the run is the same for all the runs.



**Table S1.** Number of genotyped individuals (N) and allele frequency found on locus *INRA124* in 37 bovine cattle breeds and some Zebu cattle samples.

Breed	Country	N	<i>INRA124</i> Alleles				
			126	128	130	132	134
Alistano-Sanabresa	Spain	2	0.00	0.00	1.00	0.00	0.00
Asturiana de la Montaña	Spain	6	0.00	0.00	0.58	0.42	0.00
Asturiana de los Valles	Spain	47	0.00	0.00	0.56	0.44	0.00
Avileña-Negra Ibérica	Spain	3	0.00	0.00	1.00	0.00	0.00
Blanc Bleu Belge	Belgium	3	0.00	0.00	0.67	0.33	0.00
Berrenda en colorado	Spain	4	0.00	0.00	1.00	0.00	0.00
Berrenda en negro	Spain	3	0.00	0.00	1.00	0.00	0.00
Betizu	Spain	3	0.00	0.00	0.83	0.00	0.17
Blonde d'Aquitaine	France	2	0.00	0.00	1.00	0.00	0.00
Blanca Cacereña	Spain	1	0.00	0.00	1.00	0.00	0.00
Canaria	Spain	1	0.00	0.00	1.00	0.00	0.00
Cárdena Andaluza	Spain	1	0.00	0.00	1.00	0.00	0.00
Charolais	France	2	0.00	0.00	1.00	0.00	0.00
Colombian Criollo	Colombia	1	0.00	0.00	0.50	0.50	0.00
Danish Red	Denmark	9	0.00	0.00	0.56	0.44	0.00
Dexter	Ireland	6	0.00	0.00	0.25	0.75	0.00
Fleckvieh	Austria	3	0.00	0.00	1.00	0.00	0.00
Holstein	Spain	6	0.00	0.00	0.67	0.33	0.00
Hereford	UK	5	0.00	0.00	0.50	0.40	0.10
Jersey	UK	5	0.00	0.00	0.50	0.00	0.50
Toro de Lidia	Spain	25	0.00	0.00	0.96	0.04	0.00
Limousin	France	3	0.00	0.00	1.00	0.00	0.00
Morucha	Spain	6	0.00	0.00	1.00	0.00	0.00
Normande	France	2	0.00	0.00	0.75	0.25	0.00
Pajuna	Spain	3	0.00	0.00	1.00	0.00	0.00
Palmera	Spain	1	0.00	0.00	0.50	0.50	0.00
Parla Alpina	Spain	3	0.00	0.00	0.83	0.17	0.00
Parthenais	France	2	0.00	0.00	0.75	0.00	0.25
Pirenáica	Spain	4	0.00	0.00	1.00	0.00	0.00
Retinta	Spain	5	0.00	0.00	1.00	0.00	0.00
Rubia Gallega	Spain	10	0.00	0.00	1.00	0.00	0.00
Salers	France	4	0.00	0.00	0.88	0.00	0.13
Simmental	Croatia	5	0.00	0.00	1.00	0.00	0.00
Tarantaise	France	3	0.00	0.00	1.00	0.00	0.00
Tudanca	Spain	2	0.00	0.00	0.75	0.25	0.00
Brangus	Argentina	2	0.00	0.00	0.00	1.00	0.00
Brahman	Argentina	4	0.00	1.00	0.00	0.00	0.00
Indian Zebu	India	4	0.06	0.44	0.00	0.50	0.00

**Table S1.** Description of microsatellites tested in this study. Primer sequences, reference, male specificity, size range, polymorphism, scoring repeatability and paternal compatibility are given. Tests were sequentially carried out following Pérez-Pardal et al. (2009). When a given marker did not fulfil one of our requirements no more proofs were done for this microsatellite. PCR was performed in a GenAmp PCR system 9700 (Applied Biosystems, Madrid, Spain). The final reaction volume was 10 µl. The reaction mixture consisted on 50-100 ng. of genomic DNA, 0.2 units of Taq polymerase (Biotools, Madrid, Spain), and a final concentration of 0.2 µM of each primer (Sigma-Aldrich, Madrid, Spain), 200 µM of each dNTP (Biotools, Madrid, Spain) and 2 mM MgCl<sub>2</sub>. The PCR protocol included an initial step of 95 °C (3 min), followed by 30 cycles of 30 s at 95 °C for DNA denaturation, 30 s for primer annealing at 58, 60 and 62 °C and 30 s at 72 °C for primer extension. An additional final step of 72 °C for 90 min was included to avoid n-1 band presence when using capillary electrophoresis. (Liu, Mariani et al. 2002). Parental compatibility was tested on different Asturiana de los Valles cattle breed families (Pérez-Pardal et al. 2009). Families were genotyped three times for all markers using different PCR amplification conditions, including two different taq polymerase enzymes (Biotools and TaqGold - Applied Biosystems-), to verify consistency of results. Sequentially, the markers were tested for: a) amplification on female DNA; b) amplification of a consistent number of bands (Liu et al. 2003); c) polymorphism; and d) paternal compatibility.

Marker	Primer UP	Primer DOWN	Reference	PCR conditions (°C)	Amplification on female DNA	Size range (pb) <sup>2</sup>	Polymorphism	Paternal compatibility
UMN1514	CTTCTCGAGAGTGTTCAGT	TATTCACAAGGCTCTGGAC	Liu et al. 2002	58	yes*	IMM <sup>3</sup>	-	-
UMN0910	AACAGAGATCTTTGAGCA	CTGTTCTCAGAGACTGTG	Liu et al. 2002	58	yes	IMM	-	-
UMN2001	TCAGGCAAGACTACTGGAGC	TACCTGGGCAATCTGCAA	Liu et al. 2002	58	yes	IMM	-	-
Idvga	GCATTGTATTCATACTTTCTTA	CTCTCACC TTCCGTGTGTCACT	Mezzelani et al. 1995	58	yes	IMM	-	-
DBY-1	ATCCTATAGATAGATTTAAG	TGTGGTATCTGTAATTATAC	Gotherstrom et al. 2005	58	yes	124   214	-	-
UMN2404	GGTACAATTGAAATATG	TGTACCTACACTGATATGTT	Liu et al. 2002	58	yes	214	-	-
UMN0311	GTGAGGTCTCTTGCATCTG	CCAACAATGCTTCATCCTC	Liu et al. 2002	58	no	IMM	-	-
UMN0406	GTGAGGACTCTTGCATCTG	TGCTTCATCTTCATCCAC	Liu et al. 2002	58	no	IMM	-	-
dyz	CTGGCGACTGTGCATATTT	CACACACACACCGGTTCT	Bradley et al. 1994	no amplification	-	-	-	-
UMN1203	AACAGTTCGGCACTCACCA	AGGCGACTTGTTCACAAGT	Liu et al. 2002	no amplification	-	-	-	-
UMN0705	TATAGCTGAGACGCTGAGT	CTCCTCCAGACCTGTGTATG	Liu et al. 2002	58	yes	169   189   195	no	-
UMN0504	AGGCCATCTGCATAGTGAAG	TGCTGGAGTCTCATCTCTG	Liu et al. 2002	58	yes	106   144	no	-
UMN0907	CTGTTGATCTTCTTCTCG	CTGATGGACATCTGATATC	Liu et al. 2002	55	yes	147	no	-
UMN2008	CAAGCATATCAGTGGCCCTGG	GCTGCAAGGAACTATTCA	Liu et al. 2002	58	yes	134	no	-
UMN1201	TCACCTTCCTGTGTCTATCTG	GAACAAGACCGCATTGCCA	Liu et al. 2002	58	no	IMM	no	-
INRA057	CCTAGCAGCTGTCACGCG	CACGGCTGAGAATCAAC	Vaiman et al. 1994	55	no	IMM	no	-
UMN1113	ACAGCAGCTTCTTAACAAGC	TAGCCACACATCATGTTC	Liu et al. 2002	58	no	122   134	no	-
UMN1307	GAGTTGATCTTGTGGCAG	ACATGTCAAGGTCACACAC	Liu et al. 2002	58	no	190	no	-

UMN1308	TTGTGAGAGTATTTCATGG	TCTGGACTCGACAGGACAC	Liu et al. 2002	58		no	177	no	-
UMN1605	CCACACTGAACCTGCCTCAACT	AACTATCTGCTAAGGATTGG	Liu et al. 2002	55		no	174	no	-
UMN2811	CCAGACTTGTGTATGCTGAG	CTAAAATAACTGACATGGG	Liu et al. 2002	55		no	160	no	-
UMN2905	CAGATAAGATGCGAAGTC	GCAAGTGTGACGTTAAACC	Liu et al. 2002	58		no	116	no	-
SRM18	GGCATCCAAACAGGATCAGCAAT	GTGATGGCAGTTCACATCTCTCT	Meadows et al., 2006	58		no	106	no	-
UMN0301	GCCTTGCTAGTGGCGCAACC	CAAAAGCTGTGCATGTTTC	Liu et al. 2002	58		yes*	104	no	-
UMN2303	TACTTGCTTGAGACTTACTG	TGTGAACATCTGATTTCTG	Liu et al. 2002	58		yes	IMM	yes	yes
UMN0803	TGATATTCACAGCCCGCTG	GGCTGTGGTATATCTATGAG	Liu et al. 2002	58		yes	260	yes	-
UMN0803	GATCCATGCCCGCTCAC	CTGCTTCTCTTGTCCGCTAA	Liu et al. 2002	58		yes	192	yes	yes
UMN0108	GATCATCCACATTTGCTCCA	CCAAAGCTCCATCAATTTAC	Liu et al. 2002	58		yes	203	yes	-
UMN2908	GGACTGAAGCGAGTTAGCAC	CACATCCCTGCTCACACAG	Liu et al. 2002	55		yes	188	yes	-
INRA126	GTTTGGCTCTGCAGAGTAGG	GACACTCTTCTATTTTCAAGG	Vaiman et al. 1994	58		yes	(181-187)	yes	no
UMN0929	ACCAAGCTGATACACAAAGTC	GGTCAGAGAATGAACAGAG	Liu et al. 2002	58		yes	174	yes	no
UMN0905	ATCAACCGTGTAGCTCTAA	CTAGAATGTAAACCCAGCTGC	Liu et al. 2002	58		yes	(158-168)	yes	yes
INRA124	GATCTTTGCACTGGTTTG	TTGTTAGGACTCTTGATC	Vaiman et al. 1994	55		yes	(126-131)	yes	yes
UMN0920	GTTGAGGACTCTTGATCTG	CACAGCCTAGAAGATTGAG	Liu et al. 2002	58		no	IMM	yes	yes
UMN2405	CCTGCCATCCATTGTGAAGA	CTGCTTACC TGGTCAGGATT	Liu et al. 2002	58		no	IMM	yes	yes
UMN2102	TCACCTTCTCTGTGTCATCTG	TCACCTTCTCTGTGTCATCTG	Liu et al. 2002	58		no	IMM	yes	yes
UMN2713	GTACCTACACTAATGTGTA	CCAAAGAAAGTTACAGTACA	Liu et al. 2002	58		no	IMM	yes	yes
INRA082	TCACGTGACAAACACCTTCAGAC	TGTACACATGCTGCTGTGTC	Vaiman et al. 1994	58		no	IMM	yes	-
UMN2706	TTGTTGAGGACTCTTGATC	CCACATATCAGGCAAGTCAT	Liu et al. 2002	58		no	IMM	yes	yes
BYM-1	CCTTGTTGAGCTTGACCACT	TTGCAGGCACAGAAACGGA	Mathews and Read 1991	60		no	(248-254)	yes	yes
BM861	TTGAGCCACCTGGAAGC	CAAGCGTTGGTTCAGATG	Bishop et al. 1994	62		no	(154-158)	yes	yes
INRA189	TTTTTTTCCCGTGCTGAG	GAACCTCTCTCTTGTAGCC	Kappes et al. 1997	58		no	(146-164)	yes	yes
UMN0307	GATACAGCTGAGTACTAAC	GTGCAGACATCTGAGCTGTG	Liu et al. 2002	58		no	(136-156)	yes	yes
UMN0103	ACACAGAGTATTCACCTGAG	ATTACCTGGGTCAAAGCAC	Liu et al. 2002	58		no	(112-130)	yes	yes

<sup>1</sup>Y-specific marker amplifies different product from both male and female DNA.

<sup>2</sup>Size range is given in brackets. Markers showing with 2 or 3 bands are given in different box.

<sup>3</sup>IMM: Interspersed Multilocus Microsatellites. IMM are those loci yielding several amplified bands differing in size, from the same male individual, using a single PCR reaction performed with a single primer pair. These loci are assumed to be interspersed and result from the repetitive nature of the Y-chromosome

## References



- Bishop M.D., Kappes S.M., Keele J.W., Stone R.T., Sunden S.L., Hawkins G.A., Toldo S.S., Fries R., Grosz M.D., Yoo J., Beattie C.W. (1994). A genetic linkage map for cattle. *Genetics* **136**, 619-639.
- Liu W.S., Mariani P., Beattie C.W., Alexander L.J., Ponce de León F.A. (2002) A radiation hybrid map for the bovine Y Chromosome. *Mammalian Genome* **13**: 320-326.
- Liu W-S., Beattie C.W., Ponce de León F.A. (2003). Bovine Y chromosome microsatellite polymorphism. *Cytogenetic and Genome Research* **10**: 53-58.
- Kappes S.M., Keele J.W., Stone R.T., McGraw R.A., Sonstegard T.S., Smith T.P., Lopez-Corrales N.L., Beattie C.W. (1997). A second-generation linkage map of the bovine genome. *Genome Research* **7**: 235-249.
- Matthews M.E., Reed K.C. (1991). A DNA sequence that is present in both sexes of Artiodactyla is repeated on the Y chromosome of cattle, sheep and goats. *Cytogenetics and Cell Genetics* **56**: 40-44.
- Meadows J.R.S., Hanotte O., Drögemüller C., Calvo J., Godfrey R., Coltman D., Maddox J.F., Marzanov N., Kantanen J., Kijas J.W. (2006) Globally dispersed Y chromosomal haplotypes in wild and domestic sheep. *Animal Genetics* **37**, 444-453.
- Mezzelani, A., Y. Zhang, L. Redaelli, B. Castiglioni, P. Leone, J.L. Williams, S. Solinas-Toldo, G. Wigger, R. Fries and L. Ferretti. 1995. Chromosomal localization and molecular characterization of 53 cosmid-derived bovine microsatellites. *Mamma Genome*, **6**: 629-625.
- Pérez-Pardal L., Royo L.J., Álvarez I., Ponce de León F.A., Fernández I., Casais R., Goyache F. (2008). Female segregation patterns of the putative Y-chromosome specific microsatellite markers INRA124 and INRA126 do not support their use for cattle population studies. *Anim. Genet.* doi:10.1111/j.1365-2052.2009.01870.x
- Vaiman D., Imam-Ghali M., Moazami-Goudarzi K., Guerin G., Grohs C., Leveziel H., Saidi-Mehtar N. (1994). Conservation of a syntenic group of microsatellite loci between cattle and sheep. *Mammalian Genome* **5**, 310-314.

[illegible]



### Supplementary Online Material for

#### **Multiple paternal origins of domestic cattle revealed by Y-specific interspersed multilocus microsatellites**

Lucía Pérez-Pardal, Luis J. Royo, Albano Beja-Pereira, Shanyuan Chen, Rodolfo J. C. Cantet, Amadou Traoré, Ino Curik, Johann Sölkner, Riccardo Bozzi, Iván Fernández, Isabel Álvarez, Juan Pablo Gutiérrez, Enrique Gómez, F. Abel Ponce de León, Félix Goyache<sup>1</sup>

<sup>1</sup>To whom correspondence and requests should be addressed (e-mail: fgoyache@serida.org)

#### **This PDF file includes:**

**Supplementary Table S1:** Description of sampling.

**Supplementary Table S3:** Consistency with the haplogroups previously identified by Götherström et al. (2005).

**Supplementary Figure S1:** Detailed representation of cattle Y-chromosome clades.

**Supplementary Figure S2:** Electropherograms showing paternal compatibility of the two Y-specific IMMs used.

#### **References**

**Supplementary Table S2,** containing all the genotypes and haplotypes obtained, is provided in a separated Excel file.

**Supplementary Table S1** Distribution of the 453 male individuals analysed per continent, country of origin and breed or zone of sampling. The number of samples per haplogroup (Y1, Y2 and Y3) and breed is also detailed. Additionally, each breed is assigned to a given geographical area: a) Atlantic Europe (103 samples); b) Continental Europe (248 samples); c) Mediterranean Africa (12 samples); d) Sub-Saharan Africa (32 samples); e) India (33 samples); and f) Rest of Asia (26 samples).

Further information on sampling:

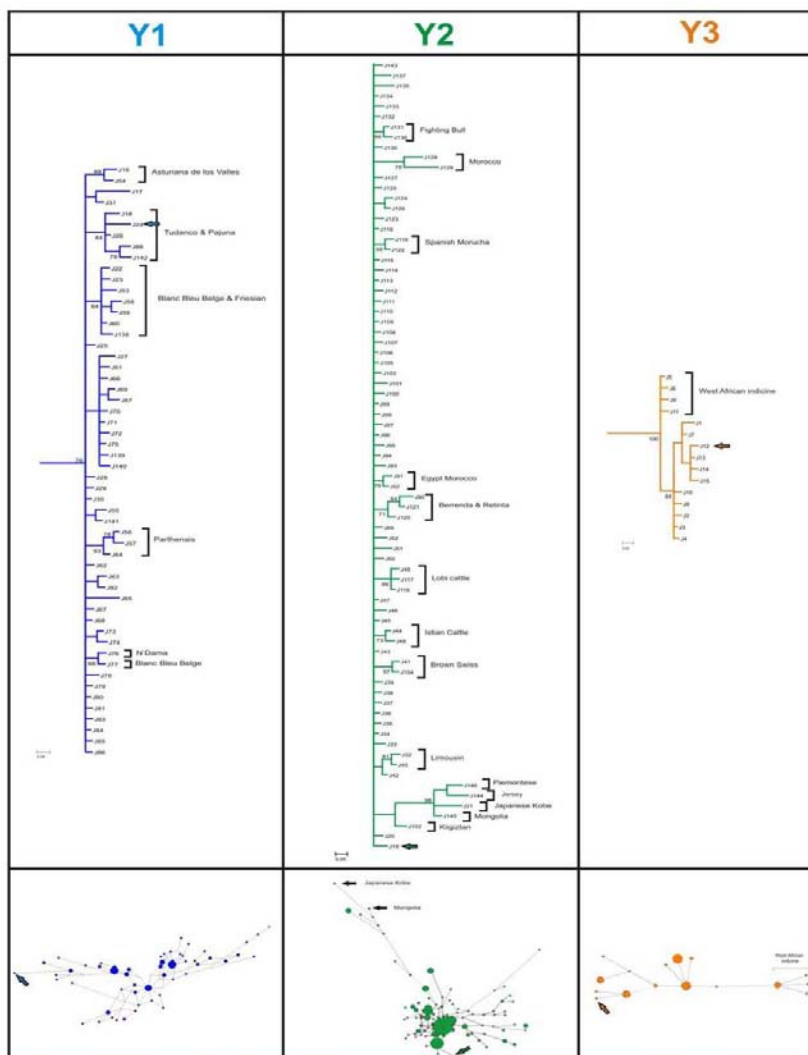
- Brahman samples were obtained from Argentinean populations. However, they are assigned to Asia.
- The samples obtained in Central India were basically sampled from the Maharashtra and Rajasthan provinces of India
- The N'Dama samples were obtained from 2 ranches of the JVL Company located in the Kolo province of Congo. The history of these herds started in 1927 with the first acquisition of 10 local bovine heads (mixed Afrikander, Baoulé, and Devon cattle) and 50 heads of Barotsé cattle from North-East Angola. From 1930, 2 bulls and 40 heifers belonging to N'dama cattle breed were directly imported from Guinea (Fouta-Djallon) to sequentially replace the more than 200 extant heads due to very poor performance. From this year replacement was selected according to performance and N'Dama phenotype. In 1945 the herd numbered 12,000 heads and consists mainly of healthy and sturdy cattle with a strong predominance of the N'dama phenotype. At present the two JVL ranches keep more than 40,000 N'Dama heads.
- The Holstein Friesian Samples mainly consisted on sires born in Spain, France, Italy and The Netherlands. However, their pedigrees could be traced back to Canada and U.S.A. To avoid inconsistencies the Holstein samples were assigned to The Netherlands, in accordance with the ancestral origin of this cattle.
- The Jersey samples were basically obtained from Belgium. A few samples from the Danish Jersey population were also analyzed.

Continent	Country		Breed area		Haplogroup frequency			Geographic Area
	Name	Frequency	Name	Frequency	Y1	Y2	Y3	
AFRICA	Burkina Faso	23	Lobi	23	0	8	15	Sub-Saharan Africa
	Congo	7	N'Dama	7	1	6	0	Sub-Saharan Africa
	Egypt	1	Egypt	1	0	1	0	Mediterranean Africa
	Morocco	11	BVOI	6	1	5	0	Mediterranean Africa
ASIA			BVTd	5	0	5	0	Mediterranean Africa
	Mozambique	2	Landim	2	2	0	0	Sub-Saharan Africa
	India	33	Brahman	7	0	0	7	India
			Central India	8	0	0	8	India
			NorthEast India	8	0	0	8	India
			Southern India	10	0	0	10	India
	Japan	1	Kobe	1	0	1	0	Rest of Asia
	Kazakhstan	5	Kazakhstan	5	2	2	1	Rest of Asia
	Kirgizstan	5	Kirgizstan	5	1	4	0	Rest of Asia

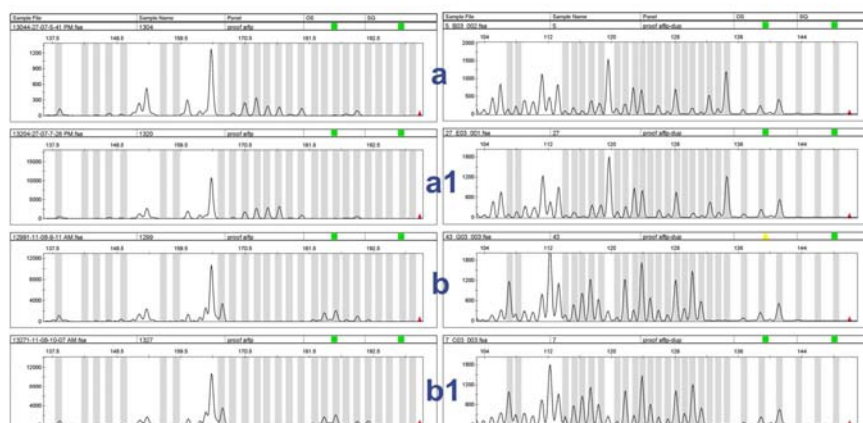
EUROPE	Mongolia	1	Mongolia	1	0	1	0	Rest of Asia
	Turkmenistan	6	Turkmenistan	6	2	2	2	Rest of Asia
	Yemen	8	Yemen	8	0	0	8	Rest of Asia
	Austria	51	Brown Swiss	15	0	15	0	Continental Europe
			Fleckvieh	6	0	6	0	Continental Europe
			Pinzgauer	9	0	9	0	Continental Europe
			Tyrolean Grey	21	0	21	0	Continental Europe
	Belgium	14	Blanc Bleu Belge	9	9	0	0	Atlantic Europe
			Jersey	5	0	5	0	Atlantic Europe
	Croatia	18	Istrian	5	0	5	0	Continental Europe
			Simmental	13	3	10	0	Continental Europe
	Denmark	6	Danish Red	6	5	1	0	Atlantic Europe
	France	64	Blonde d'Aquitaine	5	0	5	0	Continental Europe
			Charolais	15	0	15	0	Continental Europe
			Limousin	4	0	4	0	Continental Europe
			Normande	10	10	0	0	Continental Europe
			Parthenais	10	3	7	0	Continental Europe
			Salers	10	0	10	0	Continental Europe
			Tarentaise	10	0	10	0	Continental Europe
	Germany	10	European Friesian	10	10	0	0	Atlantic Europe
	The Netherlands	9	Holstein	9	9	0	0	Atlantic Europe
	Ireland	3	Dexter	3	3	0	0	Atlantic Europe
	Italy	40	Chianina	18	0	18	0	Continental Europe
			Maremmana	17	0	17	0	Continental Europe
			Piemontese	5	0	5	0	Continental Europe
	Spain	121	Asturiana de la Montaña	10	9	1	0	Atlantic Europe
			Asturiana de los Valles	29	21	8	0	Atlantic Europe
			Berrenda	5	0	5	0	Continental Europe
			Betizu	9	0	9	0	Continental Europe
			Fighting Bull	12	1	11	0	Continental Europe
			Moncha	12	0	12	0	Continental Europe
			Pajuna	4	1	3	0	Continental Europe
			Parda Alpina	8	0	8	0	Continental Europe
			Pirenaica	8	0	8	0	Continental Europe
			Retinta	6	0	6	0	Continental Europe
			Rubia Gallega	11	0	11	0	Continental Europe
			Tudanco	7	7	0	0	Atlantic Europe
	UK	15	Angus	8	8	0	0	Atlantic Europe
			Hereford	7	6	1	0	Atlantic Europe

**Supplementary Table S3** Summary of the test of the consistency between the cattle Y-chromosome haplogroups identified in this study using IMM markers with those previously identified by Götherström et al. (2005) using Y-specific loci. One sample belonging to each of the 146 haplotypes identified using IMMs was genotyped for two out of the five sites reported by these authors to avoid redundant information: zfy9 and zfy10indel. Both markers were amplified using primer pairs as described by Ginja et al. (2008) in Table 1. zfy9 amplicons were purified with Exosap (Amersham) and sequenced using Big Dye Terminator v3.1 kit (Applied Biosystems, Foster City, CA) on an ABI310 DNA sequencer (Applied Biosystems) and analyzed with Analyses 5.2 software (Applied Biosystems). zfy10indel was genotyped on an ABI3130 DNA sequencer (Applied Biosystems) and analyzed with GeneMapper software (Applied Biosystems). A summary of the obtained results is given in the table below. All samples belonging to each Y-chromosome haplogroups identified using IMMs gave the same genotype which in turn, was fully consistent with those defined by Götherström et al. (2005) as haplogroups Y1, Y2 and Y3.

Haplogroup	locus	
	zfy9	zfy10indel
Y1	C	--
Y2	C	GT
Y3	T	GT



**SupplementaryFigure S1** Detailed representation of cattle Y-chromosome clades (Y1, in blue; Y2, in green; and Y3 in orange) identified using the program MrBayes 3.1 and Network 4.11. Computations are described in the legend of the Figure 1 of the manuscript. Posterior probabilities higher than 0.60 are given to illustrate the statistical confidence on branching order. The breeds on which the haplotypes identified were found are also identified, pointing out different local diversity events. Examples of the latter are the zebuine haplotypes of the West African Lobi cattle within the Y3 cluster, the Spanish Cantabrian, Belgian Blue and European Friesian samples within the Y1 cluster and the Brown Swiss, Istrian, Southern Spanish or Mediterranean African samples within cluster Y2. Arrows without legend identify ancestral haplotypes within haplogroups according to internal number of square differences ( $d^2$ ; see Materials and methods and legend of Table 2).



**Supplementary Figure S2** Electropherograms showing paternal compatibility of the two Y-specific IMMs used for phylogenetic analyses in cattle: UMN2405, on the left; and UMN2303, on the right. Fathers are indicated with a letter and sons are identified with the corresponding letter of the father and the number 1. Grey bars represent the loci assessed as polymorphic across the whole samples used. Following Pérez-Pardal et al. (2009), amplification on female DNA and paternal compatibility was tested on two pedigrees belonging to the Asturiana de los Valles beef cattle breed. They consisted on: a) 1 sire and 3 male offspring; and b) 1 sire and 1 male offspring. Additionally, a couple of sire cousins and a couple of sire sibs belonging to two well-known Holstein grandfather families were tested (Figure S3). In all families used, paternity was confirmed with the 11 genomic StockMarks Kit for Cattle Bovine Genotyping (Applied Biosystems Inc., Foster City, CA).

## References

- Ginja C, Telo da Gama L, Penedo MCT (2008). Y Chromosome haplotype analysis in Portuguese cattle breeds using SNPs and STRs. *J. Hered.* 100: 148-157.
- Götherström A, Anderung C, Hellborg L, Elburg R, Smith C, Bradley DG et al. (2005). Cattle domestication in the Near East was followed by hybridization with aurochs bulls in Europe. *Proc. R. Soc. B* 272: 2345-2350.
- Pérez-Pardal L, Royo LJ, Álvarez I, Ponce de León FA, Fernández I, Casais R et al. (2009). Female segregation patterns of the putative Y-chromosome specific microsatellite markers INRA124 and INRA126 do not support their use for cattle population studies. *Anim. Genet.* 40: 560-564.



**Tabla S1** URLs used for sequence comparisons were: BLAST, <http://www.ncbi.nlm.nih.gov/genome/seq/page.cgi?FpHsBlast.html&pHs> (for comparing sequences with the cow genome sequence); GenBank, <http://www.ncbi.nlm.nih.gov/Genbank/> (for sequence information of the cow Y chromosome); and RepeatMasker Web Server, <http://www.repeat-masker.org> (for identifying MEs adjacent to microsatellites). Clones AC172950.4, AC226431.3 and AC167351.3 were analyzed using the RepeatMasker program (from 14000 to 25000; 1 to 7000 and 63000 to 74000 respectively). In this table we show the window obtained for Clone AC167351.3 (from 65000 to 75000). Repetition of microsatellite UMN0103 is in position 3901 to 3946.

±	score	%	div.	%	del.	%	ins.	query sequence	position in query-			C matching + repeat	repeat class/family	-position in repeat-			linkage id/graphic
									begin	end	(left)			(left) begin	end	begin (left)	
±	626	20.6	8.4	1.1	0.3	0.3	0.3	UnnamedSequence	15	181	(9820)	+ MLTIF1-int	LTR/ERV1-MaLR	1125	1303	(186)	1
±	1813	18.8	2.3	0.3	0.3	0.3	0.3	UnnamedSequence	189	540	(9461)	+ ART2A	SINE/RTE-BovB	165	523	(28)	2
±	301	28.7	2.2	2.2	0.3	0.3	0.3	UnnamedSequence	541	629	(9372)	+ MLTIF1-int	LTR/ERV1-MaLR	1315	1403	(86)	1
±	1806	19.1	16.1	0.9	0.3	0.3	0.3	UnnamedSequence	709	1199	(8802)	+ MLTIF1	LTR/ERV1-MaLR	5	569	(0)	1
±	232	20.8	3.8	0.0	0.3	0.3	0.3	UnnamedSequence	1253	1305	(8696)	+ MLTIF2	LTR/ERV1-MaLR	506	560	(0)	3
±	645	21.7	8.0	7.6	0.3	0.3	0.3	UnnamedSequence	1344	1567	(8434)	+ MER31A	LTR/ERV1	89	315	(170)	4
±	489	22.7	12.2	1.2	0.3	0.3	0.3	UnnamedSequence	1615	1770	(8231)	+ MER31A	LTR/ERV1	305	477	(8)	4
±	287	16.7	1.1	4.7	0.3	0.3	0.3	UnnamedSequence	2659	2746	(7255)	+ A-rich	Low_complexity	1	85	(0)	5
±	396	0.0	0.0	0.0	0.3	0.3	0.3	UnnamedSequence	3314	3357	(6644)	+ (TG)n	Simple_repeat	1	44	(0)	6
±	343	34.6	6.6	1.7	0.3	0.3	0.3	UnnamedSequence	3392	3655	(6346)	C L2a	LINE/L2	(168)	3207	2933	7
±	414	0.0	0.0	0.0	0.3	0.3	0.3	UnnamedSequence	3901	3946	(6055)	+ (CA)n	Simple_repeat	2	47	(0)	8
±	715	26.4	0.0	1.0	0.3	0.3	0.3	UnnamedSequence	7351	7560	(2441)	+ YREP_BOV	Unknown	1	208	(13)	9
±	229	32.1	3.6	0.0	0.3	0.3	0.3	UnnamedSequence	7700	7783	(2218)	C MIRb	SINE/MIR	(75)	193	107	10
±	22	59.1	0.0	0.0	0.3	0.3	0.3	UnnamedSequence	7860	7881	(2120)	+ AT-rich	Low_complexity	1	22	(0)	11
±	288	37.9	4.5	0.7	0.3	0.3	0.3	UnnamedSequence	7946	8078	(1923)	+ MIR	SINE/MIR	2	139	(123)	12
±	2345	20.4	9.2	5.8	0.3	0.3	0.3	UnnamedSequence	8475	8512	(1489)	+ LIMB4	LINE/L1	3643	3680	(2466)	13
±	234	0.0	0.0	0.0	0.3	0.3	0.3	UnnamedSequence	8513	8538	(1463)	+ (CA)n	Simple_repeat	1	26	(0)	14
±	2345	20.4	9.2	5.8	0.3	0.3	0.3	UnnamedSequence	8539	9252	(749)	+ LIMB4	LINE/L1	3681	4418	(1728)	13
±	711	25.8	2.8	3.2	0.3	0.3	0.3	UnnamedSequence	9254	9501	(500)	+ LI-2_BT	LINE/L1	1628	1874	(141)	15
±	1124	24.2	7.3	0.6	0.3	0.3	0.3	UnnamedSequence	9601	10001	(0)	+ LIMB4	LINE/L1	5484	5912	(268)	13

**Figure S1** Genmapper image of UMN0103 amplification obtained using the standard PCR conditions described in Pérez-Pardal *et al.* (2010a). Blue lines are due to FAM labelling of PCR products. Bands of different size from the GenScan 500-LIZ standard are in orange. Plot 1) represent a 122/126 animal belonging to the Y1 haplogroup. Plot 2) represent a 120/130 animal belonging to the Y2 haplogroup. Plot 3) represent an African 128-genotype belonging to the Y2 haplogroup. Plot 4) present a 114/ 120 Y3 animal. Note the differences between allele sizes in Plots 1) and 2).

